

PROGRAM EVALUATION CAPACITY FOR NONPROFIT HUMAN
SERVICES ORGANIZATIONS:
AN ANALYSIS OF DETERMINING FACTORS

Salvatore Alaimo

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Philanthropic Studies,
Indiana University

October 2008

Accepted by the Faculty of Indiana University, in partial
fulfillment of the requirements for the degree of Doctor of Philosophy.

David A. Reingold Ph.D., Chair

Debra Mesch Ph.D.

Doctoral Committee

David Van Slyke Ph.D.

Date of Defense
July 11, 2008

Patrick Rooney Ph.D.

© 2008

Salvatore Alaimo

ALL RIGHTS RESERVED

ACKNOWLEDGEMENTS

This research would not have been possible without the support of numerous people and organizations. The Center on Philanthropy deserves special recognition for having the vision and earnestness to take a risk and invest in a new interdisciplinary Ph.D. program in Philanthropic Studies at Indiana University-Purdue University Indianapolis (IUPUI), which I am a proud member of the first cohort of seven students. Robert Payton; Dr. Eugene Tempel, Executive Director; Dr. Dwight Burlingame, Associate Executive Director and Director of Academic Programs; Dr. Leslie Lenkowsky, Director of Graduate Studies; the center's administrative staff and Board of Directors; and the entire philanthropic studies faculty are recognized for getting this new program off the ground and ensuring it will be continually supported. My six cohorts have been a source of support throughout this Ph.D. program, and we continue to exchange information and moral support today. These friendships have impacted my work and my life, and I hope we will continue to be connected for years to come. They are a large part of why I feel Indianapolis is my second home.

The majority of my research expenses were self-supported; however Jill Kramer and The Lumina Foundation are recognized for contributing a generous grant to support some of these expenses. The Center on Philanthropy and Patrick Rooney, Director of Research, and the IUPUI Graduate Student Organization are also recognized for their interest in and financial support for this research.

I want to thank my dissertation committee members, Dr. David Reingold, Dr. Debra Mesch, Dr. David Van Slyke, and Dr. Patrick Rooney. Their wealth of knowledge, insight and guidance were critical to the success of this research.

Special thanks go to my cousin Thomas Alaimo for providing free flights to Indianapolis, and to Hank and Julie Niec for allowing me to stay at their home numerous times and being such gracious hosts. My former employer the Girl Scout Council of Northwest Georgia and former CEO Harriet Hessam deserve special recognition for providing a career opportunity in evaluation that sparked my interest in this topic. This research would not have been possible without the people who volunteered their time as anonymous interview participants. The case studies were possible thanks to the transparency and cooperation from Pat Showell, Chris Valley, and the staff and board at Families First; Dr. Peter Lyons from Georgia State University; Beth Vann, and the staff and board at Decatur Cooperative Ministry.

Many people in both the metro Atlanta and Indianapolis areas were responsible for providing meeting space, expertise, opportunities to present my research which resulted in valuable feedback from audiences, and for contributing to the pre-testing of interview instruments. These contributions enabled this work to be successfully completed. They include: Dr. Brian Vargus, Laura Littlepage, Fran Huels and the staff at the Joseph and Matthew Payton Philanthropic Studies Library, the staff at the IUPUI University Library, Brenda Burk at the Ruth Lilly Special Collections and Archives, the staff at IUPUI's Research Compliance Administration, the American Evaluation Association (AEA), the discussants on EvalTalk, the Association for Research on Nonprofit Organizations and Voluntary Action (ARNOVA), the Atlanta Area Evaluation Association (AAEA), the Indiana Evaluation Association (IEA), the American Association for Public Opinion (AAPOR), the Academy of Management, GuideStar, Jan Diggins at Citizens Gas, Sarah Cole at Emmis Communications, Robert Eger at

Children's Restoration Network, Rhett Tanner at Senior Citizen Services of Metropolitan Atlanta, Donovan Quarry of the Make a Difference Foundation, Carlton Sapp at Gwinnett County Public Schools, Pattie Johnson and the Foundation Center staff in Atlanta, Wesley Simms at the Indiana Grantmakers Alliance, Alice McColgin at Indianapolis Power & Light Company, Kendy Smith at Eli Lilly and Company, Jill Wohlford at the Lumina Foundation, Janice Robinson at the United Way of Metropolitan Atlanta, Boris Volkov and Jean King from the University of Minnesota, Lynne Ensminger at the Blue River Community Foundation, Joanne Carman from the University of North Carolina-Charlotte, Kim Fredericks from Indiana State University, Deena Murphy-Medley at National Development and Research Institutes, Inc., Don Compton, Tom Chapel, Maureen Wilce and Kai Young at the Centers for Disease Control, Jeff Brudney from Cleveland State University (formerly from the University of Georgia), Al Roberts from Syracuse University, and Dennis Young from Georgia State University. Special thanks go to Ting-Ting Liang, Molly Neu, Jennifer Wewers, Telesa Wilson, Dorothy Goodwin, Alice Meyers, Tangela Love and Judy Headrick for their assistance with transcribing tapes of the interviews.

I appreciate the references and sources for this research, as well as continual words of encouragement from my lifelong friend Steve Owens. My parents, Thomas and Angie Alaimo, continue to provide the strong support and encouragement they have provided all of my life. Their support has not only contributed to the success of this research, it also has contributed to who I am as a person. I hope to retain these values the rest of my life, as I continue to grow as a person and expand my intellectual endeavors. I continue to thank my life partner Lee Ann Harvey for her ongoing support throughout the

tenure of this research, especially during the first two years of separation for the PhD program's coursework. In spite of many hours spent on this research and away from her, she encouraged me to focus, prioritize my time and always keep the end in sight.

ABSTRACT

Salvatore Alaimo

PROGRAM EVALUATION CAPACITY FOR NONPROFIT HUMAN SERVICES ORGANIZATIONS: AN ANALYSIS OF DETERMINING FACTORS

The increasing call for accountability combined with increasing competition for resources has given program evaluation more importance, prominence and attention within the United States nonprofit sector. It has become a major focus for nonprofit leaders, funders, accrediting organizations, board members, individual donors, the media and scholars. Within this focus however there is emerging attention and literature on the concept of evaluation capacity building to discover what organizations require to be able to effectively and efficiently evaluate their programs.

This study examines this topic within the environment and stakeholder relationship dynamics of nonprofit human service organizations. A multi-stakeholder research approach using qualitative interviews of executive directors, board chairs, program staff, funders and evaluators, as well as two case studies, is employed to provide insight into the factors that determine an organization's evaluation capacity. The overarching goal of this research is to impart this information to stakeholders interested in program evaluation, by analyzing elements for capacity beyond the more common, narrow scope of financial resources and evaluation skills. This purposeful approach intends to broaden our understanding of evaluation capacity building to encompass developing the necessary resources, culture, leadership and environments in which meaningful evaluations can be conducted for nonprofit human service programs.

Results indicated that effective evaluation capacity building requires more than just funds, personnel and expertise. Some of the important factors that impacted this process included leadership; value orientations; congruence among stakeholders for their perceptions of evaluation terms and concepts; resource dependency; quality signaling; stakeholder involvement and understanding of their role in program evaluation; organizational culture; organizational learning; personal preferences; and the utilization of available evaluation tools. This study suggests that stakeholders interested in effectively building capacity to evaluate programs should be cognizant of these political, financial, social, intellectual, practical, structural, cultural and contextual implications.

David A. Reingold Ph.D., Chair

TABLE OF CONTENTS

List of Tables	xiii
List of Figures	xiv
Chapter One	
The Importance of Studying Program Evaluation.....	1
Significance of the Study	1
Statement of the Problem to be Studied	3
Chapter Two	
Review of Related Literature.....	6
Formal Organizations.....	6
Human Services	33
Program Evaluation.....	39
Evaluation Capacity	56
Constructs	56
Approaches	73
Nonprofit Management	86
The Role of Mission.....	87
NHSOs and Program Evaluation	89
Organizational Culture	103
Leadership and Organizational Culture	112
Managing the Socio-Political Environment	117
Meeting the Challenges for Program Evaluation	125
Courses of Action for Program Evaluation.....	128

Organizational Learning.....	132
Stakeholder Roles.....	136
Executive Directors	138
Board of Directors	142
Program Staff.....	144
Funders	151
Evaluators	157
Chapter Three	
Methodology.....	170
Research Questions	170
Sampling	170
Sampling Procedures.....	172
Sampling Procedures by Stakeholder Group	175
Instruments.....	177
Data Collection.....	192
Data Analysis	192
Limitations of Study.....	194
Chapter Four	
Results	202
NHSOs	202
Executive Directors	206
Board Chairs.....	213
Program Staff	224

Funders	240
Evaluators.....	248
Case Studies	260
Families First	260
Decatur Cooperative Ministry	269
Chapter Five	
Conclusions and Recommendations	275
Chapter Six	
Implications for Practice and Further Research	288
Appendices	
Appendix A.....	293
Appendix B	294
Appendix C	295
Appendix D.....	298
Appendix E	303
Appendix F.....	308
Appendix G.....	311
References.....	314
Notes	331
Curriculum Vitae	

LIST OF TABLES

Tables

Table 1	Common Principles and Standards among Professions	55
Table 2	Cultural Products for Program Evaluation by Organizational Variable	109
Table 3	Funder Roles in Empowerment Evaluation	155
Table 4	CIPP Model for Evaluation	156
Table 5	Sampling of Stakeholder Groups	171
Table 6	Organization Rating Types for Program Evaluation Activity.....	204
Table 7	Organizational Characteristics by Type	205
Table 8	Common ED Perceptions for Program Evaluation	207
Table 9	ED Profile Information by Organization Type	212
Table 10	Board Chairs' Descriptions of Accountability	216
Table 11	Top 3 Priorities for the Board and Board Chair	218
Table 12	Board Chair Profile Information by Organization Type	223
Table 13	Program Staff Profile Information by Organization Type	240
Table 14	Funder Profile Information	247
Table 15	Evaluator Profile Information	259

LIST OF FIGURES

Figures

Figure 1	Stakeholders Invested in Program Evaluation	3
Figure 2	Balancing Fiscal and Programmatic Accountability	19
Figure 3	Elements for Evaluation Capacity Building.....	61
Figure 4	Relationship Between Mission and Program Evaluation.....	88
Figure 5	Program Improvement Continuum	91
Figure 6	Comparison of Sample to Population by County – Atlanta MSA	173
Figure 7	Comparison of Sample to Population by County – Indianapolis MSA	173
Figure 8	Comparison of Sample to Population – All NHSOs by NTEE Code	174
Figure 9	Organization Total Annual Revenue.....	203
Figure 10	Age of Organizations	203
Figure 11	Organization Type by Program Evaluation Activity.....	205
Figure 12	Executive Director Actions for Program Evaluation	207
Figure 13	Descriptions of Board’s Role in Program Evaluation.....	220
Figure 14	Board Chair’s Satisfaction with Organization’s Level of Engagement in P.E.	221
Figure 15	Frequency of Board Meetings.....	223
Figure 16	EDs Having a Role in the Program Evaluation Process	231
Figure 17	Trade-Off Costs for Program Staff by Organization Type	233
Figure 18	Information Funders ask of Grantees	242
Figure 19	Program Evaluation Information Funders ask of Grantees.....	243
Figure 20	Levels at which Funders Factor Information into Continuing to Fund.....	244

Figure 21	Funders Financially Supporting Program Evaluation	245
Figure 22	Opinion for Similar Funders' Support of Program Evaluation	246
Figure 23	Most Important Stakeholders for ECB (According to Evaluators)	254
Figure 24	Considerations and Implications for ECB	292

The Importance of Studying Program Evaluation

The increasing call for accountability combined with increasing competition for resources has given program evaluation more importance, prominence and attention within the United States nonprofit sector. It has become a major focus for nonprofit leaders, funders, accrediting organizations, board members, individual donors, the media and scholars (Brody 2002; Reamer, 1998; Salamon, 2002). Some indicate that the demand for program evaluation is growing (Carman, 2007; Fitzpatrick, Sanders & Worthen, 2004; Hudson, 2005; Newcomer, Hatry & Wholey; 2004), and that it is increasingly being used in the nonprofit sector as a component of organizational performance measurement (Patton, 1997; Poister, 2003). Funders especially have suggested that if nonprofits were to more actively engage in program evaluation, their accountability to a myriad of stakeholders could be enhanced and organizational transparency could be increased (Hudson, 2005). Nonprofit organizations face the challenge of responding to the *external pull* from funders, government agencies and accrediting bodies while developing an intrinsically motivated *internal push* to build long-term capacity to evaluate their programs.

Significance of the Study

This study examines nonprofit human service organizations (NHSOs) because they particularly operate in an environment of increasing pressure for demonstrating program effectiveness resulting from the devolution of social services from government to the nonprofit sector, and more sophisticated evaluative information required by funders and accreditation requirements. NHSOs interested in formally evaluating their programs will need to comprehend what is required for their capacity to effectively and efficiently

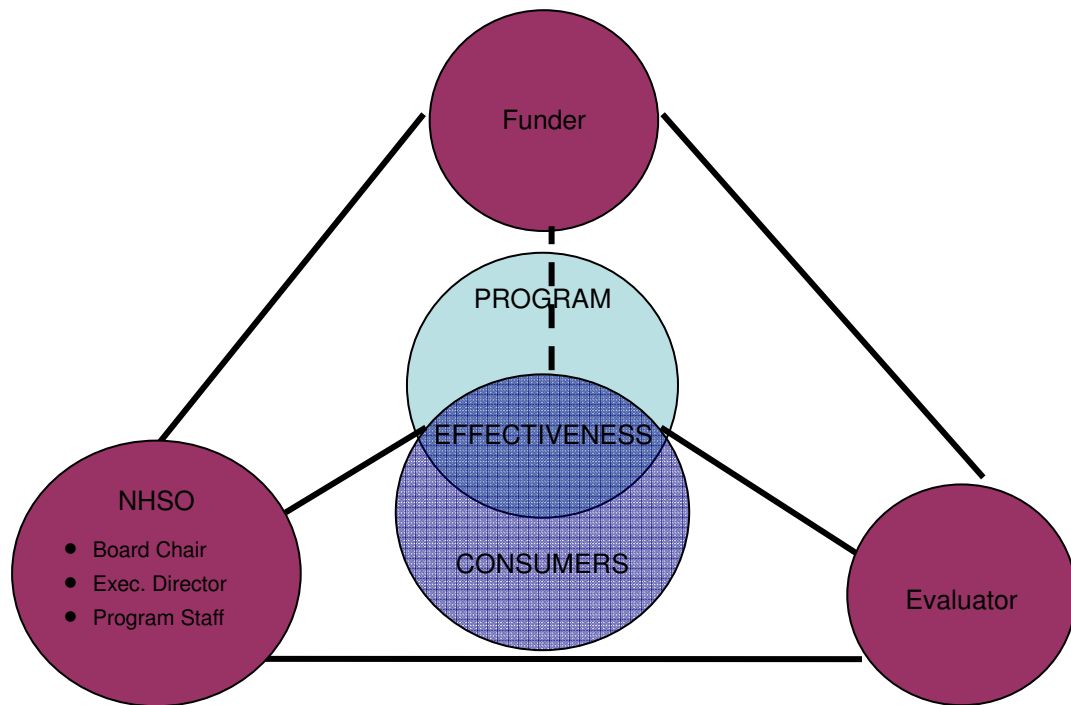
engage in the process. The literature on evaluation capacity building (ECB) emerged in the late 1990's and is still a relatively new and expanding topic. It is growing within the evaluation profession and is only recently permeating the U.S. nonprofit sector and the field of nonprofit management. As a result, there have not been many studies conducted on this topic.

The objectives of this research include providing insight into the factors that determine an organization's evaluation capacity; expanding and improving our understanding of evaluation capacity within the arena of NHSOs; contributing to the growing literature; and hopefully inspiring future, additional studies on ECB. The overarching goal of this research is to impart useful information to stakeholders interested in program evaluation, so they will have a more comprehensive understanding of what is required to build capacity for it. In order to accomplish this research goal and subsequent objectives, this study will examine ECB by 1) analyzing capacity components beyond the more common narrow scope of financial resources and evaluation skills, and 2) incorporating a multi-stakeholder approach that provides different perspectives and a more holistic approach to the topic. This purposeful approach intends to broaden our understanding of ECB to encompass developing the necessary resources, culture, leadership and environments in which meaningful evaluations can be conducted for NHSOs' programs

The three primary stakeholder groups potentially interested in or engaged in evaluating nonprofit human service programs are NHSOs, funders and evaluators as shown below in Figure 1. The specific stakeholder groups interviewed in this study

include the executive directors, board chairs and program staff of NHSOs, various types of funders, and evaluators of nonprofit human service programs.

Figure 1 – Stakeholders Invested in Program Evaluation



Statement of the Problem to be Studied

Stakeholders, such as NHSOs, funders and evaluators that have an interest in program evaluation, regardless of motivation, logically require a comprehensive understanding of what comprises this concept of evaluation capacity. An emerging literature and a few studies on evaluation capacity building have significantly contributed to this level of understanding; however the topic of ECB is still evolving. This study seeks to acquire insight on the factors that determine a NHSO's capacity to evaluate their program(s). It also attempts to examine the relationship between capacity and organizations' decisions on whether to engage in evaluating their programs. Evaluation's presence in the nonprofit management literature has grown in the last decade

(Braverman, Constantine & Slater, eds., 2004; Gray, 1995; Hoefer, 2000; Mesch & McClelland, 2006; Paddock, 2001; Thomas, 2005; United Way of America, 1996) as it has garnered more attention in the nonprofit sector. However the topic of ECB has only recently and sparsely been discussed in the U.S. nonprofit sector.

ECB is a relatively new concept and the current literature comes almost exclusively from the evaluation field and therefore emanates from the evaluator's perspective (Boyle, Lemaire & Rist, 1999; Compton, Glover-Kudon, Smith & Avery, 2002; Milstein, Chapel, Wetterhall & Cotton, 2002; Preskill & Russ-Eft, 2005; Stockdill, Baizerman & Compton, 2002; Stufflebeam, 2002). It primarily examines ECB as an activity, or series of activities, conducted internally by the organization or in collaboration with external stakeholders, most prominently the evaluator. This literature has been ground-breaking and critical to the expanding field of evaluation and the growing interest for evaluation in the nonprofit sector.

This study seeks to draw from and contribute to this growing and evolving literature by taking an interdisciplinary look at the factors that determine evaluation capacity. The review of literature will examine formal organizations and how they respond to their environment with specific subsets of organizational culture and organizational learning. It will also draw from other bodies of literature including nonprofit management, human services, leadership and evaluation.

The question of what it takes to adequately evaluate nonprofit human service programs is just beginning to be answered by the key stakeholder groups invested in program evaluation, primarily from evaluators. This study can provide some important information that helps these groups begin to more thoroughly answer this question, and

ultimately increase and improve the conducting and usage of program evaluation while improving the efficiency and effectiveness of programs.

Review of Related Literature

Formal Organizations

An effective analysis of the factors determining evaluation capacity in NHSOs first requires an understanding of formal organizations. The following section puts this relevance and importance within the context of issues that likely will affect ECB in NHSOs.

One common method for examining formal organizations found in the literature is looking at their transactions within the context of markets and hierarchies. More specifically, it is important to examine how market and non-market organizations behave and perform. This involves the concept of *bounded rationality* which plays an important role in an individual's ability to analyze and solve problems. Bounded rationality is defined as "...rational choice that takes into account the cognitive limitations of both knowledge and computational capacity" (Simon, 1997, p. 291). Theories of bounded rationality assume we desire to attain goals and therefore use our minds to the best of our ability while factoring in our limited intellectual capacities. The theories suggest we recognize that our ability for "formulating and solving complex problems is very small" when we consider that such solutions require "objectively rational behavior" (Williamson, 1975, p. 9).

Individuals have physical and language limits, to name a few, which contribute to this limited ability to address complex problems. This bounded rationality combines with uncertainty to set up economic problems. The limits of bounded rationality impede an organization's ability to deal with the environment's uncertainty unless their key social actors choose viable alternatives. We can conclude that cognitive limits determine a

NHSO's stakeholders' ability to generate and evaluate alternatives in their decision-making process. They may seek and develop strategies for dealing with their environment of uncertainty which most likely will not include the known probability of outcomes.

We can see how this might apply to situations where NHSOs have considered evaluating their programs for the purpose of addressing the uncertainty of the programs' actual impact on their consumers. The bounded rationality of not having the expertise in-house to conduct a meaningful and accurate evaluation might drive the organization to take one of several directions. They may seek that expertise from external sources and use their own resources to pay for it. They may seek external funding to pay for the external expertise. Or the lack of in-house expertise may limit them or drive them to not pursue formally evaluating their program. Another possibility is they may engage in satisficing, or the act of looking "...for a satisfactory, rather than optimal alternative" which is typically a "...course of action satisfying a number of constraints" and "...far easier to discover than a course of action maximizing some function" (Simon, 1997, p. 293). This might describe when a leader of a NHSO determines that program evaluation may be too difficult and/or too expensive to optimize internally, so the course of action is to satisfy the demands of external stakeholders in order to acquire funds to alleviate the constraints of resource dependency, satisfy the requirements of a government contract, or attain accreditation status.

Individuals manage these transactions through "their own observations on how events are changing" by "inferring probable consequences and acting accordingly" (Williamson, 1975, p. 25). This can describe how key stakeholders within a NHSO feel

they know their program more intimately than anyone else and can see its impact before their very eyes. Program staff may have developed their organization's program and in some cases may have also been delivering it to their consumers for years. Such intimate and sometimes passionate attachment to the program can contribute to their reality construct of the program. One ramification of this can be a disconnect, or explicit disagreement on program outcomes, between a NHSO and its funder or an evaluator. An evaluator that comes from the funder's perspective may see the program differently through the lens of that funder's agenda based on their funding focus areas and priorities.

Another important issue that comprises these contractual transactions is *information impactedness*. When one organization has an advantage in owning information, in some cases exclusively, the other organization involved in the contractual transaction will pay the cost in attempt to acquire that information and bring some parity to the transaction and the relationship (Williamson, 1975, p. 14). Information impactedness plays a role in dealing with information asymmetry typically found within the parameters of program evaluation. There are several scenarios that can play out within these parameters. One is where the funder is seeking the information about the program that the NHSO seemingly has. By funding the program and its evaluation, the funding organization can attempt to move the shared information about the program towards parity. Another example is when program evaluation helps bring parity to NHSOs with similar service delivery types, where one organization has endeavored in research and evaluation to where they are seen as the expert or model for that specific service delivery type. By utilizing research and evaluating its programs, an NHSO can close the gap in information asymmetry.

The level of funding, comprehensiveness of the evaluation, and level of detail in the reporting all help determine the effectiveness of this movement towards information parity. The funder may be paying a considerable cost to gain this information, but they may still ultimately be relying on the NHSO to candidly disclose the information.

Another possible scenario played out between the NHSO and its funder with respect to information impactedness needs to be considered. If the NHSO chooses to be transparent and open about its program, and the funder is sincerely interested in learning about it and investing albeit financially in its impact, the relationship can become more of a partnership where shared information and organizational learning is encouraged, developed and sustained.

Rating systems can be used in an attempt to gain parity and reduce associated risk between organizations and the information they rely on to exist and succeed (Williamson, 1975). Rating systems for nonprofits are a relatively new concept, as several charity “watchdog” organizations such as Charity Navigator¹ and the Better Business Bureau’s Wise Giving Alliance² have cropped up in attempt to monitor organizations against their established standards. However, the majority of such standards address financial matters and few deal with matters of programmatic effectiveness within overall organizational accountability. In contrast to products and services in the for-profit sector, a *Consumer Reports* for NHSOs’ programs does not exist. The long standing strategy of quality signaling that nonprofits use to attract resources remains pervasive. The point about performance audits that Williamson makes is more relative to the issue of program evaluation, as he states, “...without a performance audit, the true explanation for outcomes that are jointly dependent on the state of nature...cannot be accurately

established” (Williamson, 1975, p. 16). Simply put, the intended outcomes for a program in which the NHSO and the funder are jointly invested in cannot be determined or explained without program evaluation.

Opportunism is described by Williamson as “the lack of candor or honesty in transactions to include self-interest” (Williamson, 1975, p. 9). It can play into program evaluation when NHSOs manipulate information or misrepresent their programs to external or internal stakeholders. This can result from internal evaluations where they exercise control or it can result from manipulation of results from an external evaluation. Motivations for these choices vary from just wanting to appear that they conduct evaluations with no plan on using the information, to not wanting to look bad to the very same funders they rely on for program resources. We see clearly the power of resource dependency as a driving force behind these contractual decisions. Leaders of NHSOs may also translate any “negative” information about the program as a personal indictment of their management and/or leadership whether justified or not. Similarly, opportunism can take the form of avoidance of evaluation as a priority when the focus is on growth and expansion to serve the leader’s desire for personal gain in salary and prestige. This can come at the expense of the mission of the organization and its ability to demonstrate the outcomes of its program.

While the pressure is on the nonprofit sector for more accountability through demonstrated results and fiscal responsibility, the sector still remains largely self-regulated. One can take a cynical stance on self-regulation if it is believed that the social actors who agree to such terms do so in a casual, self-disbelieving manner. Two strategies that can effectively deal with intentions for self-regulation include spelling out

a contract's terms specifically and completely, and monitoring progress to ensure no shortcuts are taken (Williamson, 1975). Here is where, ironically, funders and accreditation bodies can actually contribute to the potential hazards of agreements with NHSOs through vague language, loose or nonexistent monitoring, and/or focusing more on how the money was managed rather than the impact of the program. Such oversight takes exception to the claim that competition from large numbers of organizations "renders opportunistic inclinations ineffectual" and that "strategic posturing is nonviable behavior" (Williamson, 1975, p. 27). This can hold true in the for-profit arena where competition tightens up contractual agreements, monitoring and enforcement among the players. However, in the funding world of the nonprofit arena where competition certainly exists, if funders that control the resources do not tighten up such contractual parameters, the environment for opportunistic behavior may be perpetuated.

The differences between an internal and external auditor may apply to situations where a NHSO may hire an internal or external evaluator to examine their program. We're reminded that external auditors are typically constrained by whatever records are available to review or by a limited ability to investigate important matters. An internal auditor, however, has more freedom and can be less formal in providing evidence (Williamson, 1975). I would add that an external evaluator may have many clients and therefore be also constrained by time limits, while an internal evaluator's only client is his place of employment and only focus their program. Some conclude that an internal evaluator (in this case auditor) will act in the interests of their organization (Williamson, 1975), while others suggest that internal evaluators (as well as external evaluators) can

utilize the Joint Standards³ and Guiding Principles⁴ as means to remain as objective as possible and deal with the inherently political context of evaluation (Stake, 2001).

To better understand what is behind the actions of an internal evaluator, it is important to understand what the interests of the NHSO are. For example, an internal evaluator may be told that no matter what the results are, only positive information that reflects well on the organization is expected to be conveyed. He or she may also be forced to conduct the evaluation within budget constraints. This may not allow for a meaningful, relevant or accurate evaluation. Another possibility is for the evaluator to be told to conduct the evaluation according to the standards and guiding principles and report on the program honestly and accurately, even if the program failed to attain its intended outcomes. Lastly, an internal evaluator may be a free-spirited entrepreneur at his or her own employment peril and decide to take risks by conducting the evaluation in defiance of the instructions he or she has received from their organization. The issues of ethics, work styles, incentives, personal integrity and organizational culture all come into play in determining what course to take. They are not limited to internal evaluators as external evaluators face them as well. Some suggest that the independence of external evaluators is a misperception because the organization contracting with them ultimately controls the administration of the evaluation (Reingold, 2006). The issue of objectivity within the political context of evaluation remains a challenge for the profession and a continuing area of discussion for the American Evaluation Association (AEA).⁵

The “new institutionalism” reminds us that organizations are open systems (DiMaggio & Powell, 1991). Consequently, there are environmental and contextual factors that drive the choices of rational actors within organizations. The new

institutionalism helps set the stage to examine what is behind the capacity for program evaluation in NHSOs – the examination of organizations, the social context in which they exist and the sociological drivers of individual behavior with respect to the possible *internal push* and/or *external pull* from their environment and stakeholders.

Institutional economists argue that individuals within institutions seek to maximize an environment of stability and consistency in which their behavior can reside, but within limited cognition, incomplete information and the difficulty of managing agreements through monitoring and enforcing them. Organizational economists feel that institutions provide dependable and efficient structures for economic exchange in an attempt to reduce uncertainty. The question remains as to whether leaders of NHSOs view program evaluation as a tool that contributes to a more stable programmatic environment through reducing uncertainty within the context of the program. Do they also view it as a means for demonstrating results that can aid their organization in attaining resources in an uncertain funding environment? Or do some leaders of NHSOs view program evaluation as both?

Organizations can survive and persist when they do not conform to external demands because, “the prospective gains from altering them are outweighed by the costs of making the changes” (North in DiMaggio & Powell, 1991, p. 4). This describes how NHSOs may decide not to evaluate their programs and meet the external demands of foundations, the United Way, governmental agencies or accreditation organizations if they determine the associated costs with altering their operations and evaluating their program outweigh what they perceive to be the gain from it. If the organization’s leadership determines it is not worth it, and they can survive without doing it, they likely

will not make an effort to conduct the evaluation. Or, they might conduct a program evaluation at a level where it satisfies these external demands through symbols and signaling instead of through demonstrating the utilization of results for the improvement of the program.

The sociological branch of institutionalism contends that individuals don't choose freely among institutions, customs, social norms or legal procedures. For example, March and Olsen state that "actors associate certain actions with certain situations by rules of appropriateness" and this is done through socialization, education, on the job training, etc. (March & Olsen in DiMaggio & Powell, 1991, p.10). Such behaviors that become institutionalized are typically slower to change, as they become viewed as social norms, traditions, and/or part of the organization's history or legacy. The cliché of "we have always done it this way" might apply in this context, so we can see the potential challenges of incorporating program evaluation into the culture and ways of work of an organization. If not engaging in evaluation has become institutionalized, it will likely be a difficult transition toward prioritizing it and eventually engaging in it. It has been suggested that for program evaluation to be truly effective, it must be institutionalized into the operations and culture of an organization (Stockdill, Baizerman & Compton, 2002; Stufflebeam, 2002; Sanders, 2003; Volkov & King, 2005). The point of contention is that this is typically required for program evaluation to gain the notoriety and priority within the organization that will drive the acquisition of necessary resources and the development of the proper environment.

Such institutionalization also can help in assuring program evaluation is conducted regardless of the external demands of the environment. This can characterize

more of a financially and culturally supported *internal push* from the organization rather than a potentially reluctant cooperation with an *external pull* from other institutions or stakeholders or an “external shock” comprising a traumatic event requiring a complete change in direction for the organization. Powell and DiMaggio remind us that “fundamental change occurs under conditions in which the social arrangements that have buttressed institutions suddenly appear problematic” (DiMaggio & Powell, 1991). The consulting field characterizes this as a client’s “pain point,” a point in time where a change in behavior occurs in response to circumstances that have become intolerable. This again begs the question of whether intensifying competition for resources and increased call for accountability will drive this fundamental change towards evaluating programs to demonstrate effectiveness. It also asks if the funding world will seek to properly provide the necessary resources to conduct the very evaluations they seek of their recipients’ programs.

The resource dependence perspective looks at how organizations attempt to adjust to external demands from their environment with particular attention to those demands from whom the organization depends on for resources. The environment in which organizations exists is constantly changing and resources are not distributed evenly or regularly (Pfeffer & Salancik, 1978). This certainly describes the environment in which nonprofit organizations, particularly NHSOs, operate as their funding is uncertain and can vary greatly as public and private funding administrations, stakeholders and priorities change. The availability of funding from these sources relies greatly on economic conditions which typically are also not even or constant. As a result, NHSOs have to

adjust to meet external demands and secure funding to operate their programs and ultimately stay in business.

External constraints affect the behavior of individuals within organizations differently depending on their presence. If there are few external constraints, people tend to find intrinsic reasons that are satisfying and that justify doing the task. Conversely, if there is a larger presence of external constraints, people may not attempt to find that intrinsic value in the task and tend to justify the task based on those constraints. This helps us understand the concept of the external pull from funding organizations that require program evaluation from their recipient organizations. When leaders of NHSOs indicate that if funders didn't require evaluation they wouldn't make an effort to measure the outcomes of their programs, this substantiates that requirement of external constraints for action in lieu of intrinsic motivation.

Comparing context versus individual's impact on activities within organizations has not been adequately addressed (Pfeffer & Salancik, 1978). While the importance of context has been made clear through the post modern literature on the new institutionalism, it appears that individuals in organizations can be just as influential for determining behavior. For example, individuals also have constraints. Some are limited by the range of skills, knowledge and characteristics they can gain while achieving important positions within organizations. Some are limited by the discretion they can exercise due to accountability systems and required approval from others.

Combining these points, we can see two factors that might drive the decision of a leader of an NHSO for whether to evaluate their organization's programs. First, if a leader of a NHSO has not had any formal or informal instruction in evaluation he or she

is less likely to prioritize it within the organization due to personal preference based on their skills, knowledge and characteristics. People tend to emphasize and prioritize areas they have come from and are most familiar with. A former fund development executive will most likely put priority on the fundraising function of the organization the same manner in which a former social worker might prioritize the direct service function of the organization. Second, a leader of a NHSO would typically have to get the approval of his or her board of directors for the evaluation process and the resources necessary for it. In this case, the leader i.e. executive director may advocate for evaluating the program but may be overruled by the board if they do not realize the importance of it or feel it is worth the expenditure.

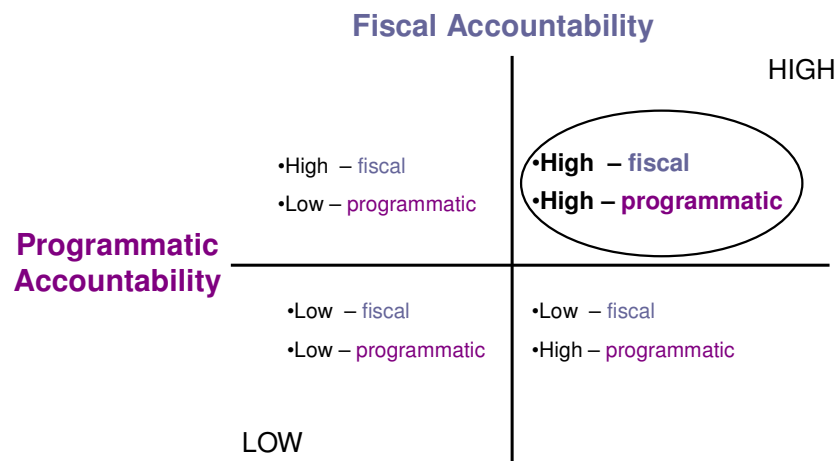
Two other factors that may play into the leader's decision making process with respect to program evaluation. One is that information used to formulate decisions comes from others and the second is that influence comes from peers (Pfeffer & Salancik, 1978). With respect to information from others, if the leader of the organization has not recognized program evaluation as an important activity or does not advocate for it, then information from others in the form of advocacy would be necessary to move it forward. Leaders of NHSOs typically belong to peer groups, usually in the form of associations, kindred services, United Way grantee groups and others. These leaders may base their decision on how far to, if at all, engage in program evaluation based on what their peers' efforts have been, an example of isomorphism. This may result regardless of what the external constraints are, as if a leader perceives the group's level of effort to become commonplace or acceptable, that may be the extent of effort he or she may engage in.

Organizational effectiveness is the ability of an organization to take actions that produce acceptable outcomes. Effectiveness is described as “an external standard of how well an organization is meeting the demands of organizations concerned with its activities” (Pfeffer & Salancik, 1978, p. 11). This description of effectiveness describes how leaders of NHSOs may view the call for program evaluation as an “external standard” from the organizations that ask for it. However, I counter that effectiveness also speaks directly to how a NHSO’s program impacts its consumers. It answers the question of whether the program has met the desired outcomes that are linked to that organization’s mission. I would also add that in this perspective, effectiveness is more of an internal standard set by the organization that develops and delivers the program. The issues of a program being effective in attaining intended outcomes and an organization meeting the external demand for evaluation from an external stakeholder should not be confused.

Effectiveness is also a sociopolitical issue that is not limited by economic factors (Pfeffer & Salancik, 1978). It also involves assessment of the usefulness of the activities and the resources required for those activities. I argue that efficiency is commonly confused, and used interchangeably, with effectiveness. A NHSO can be efficient at using resources to deliver an ineffective program that does not attain intended outcomes. Conversely, a NHSO can be inefficient with resources to deliver an effective program that attains intended outcomes. Efficiency is defined as the ratio of inputs to outputs produced and it is characterized as a managerial problem (Pfeffer & Salancik, 1978). This raises the question as to whether evaluating and improving programs is also a managerial problem. The focus on efficiency within the increasing call for accountability

is currently represented primarily by a focus on fiscal accountability i.e. efficiency and rarely includes effectiveness for programmatic accountability. Complete, overall accountability requires a balance of fiscal (efficiency) and programmatic effectiveness as illustrated in Figure 2 below.

Figure 2 – Balancing Fiscal and Programmatic Accountability



Attempting to balance these two accountabilities remains a challenge, as efficiency and effectiveness can compete and pull from each other; however the public trust that NHSOs are held in might provide incentive to strive for high accountability in both areas of operations.

The impact of an environment on an organization can vary due to buffers or filters between them or due to an organization simply recognizing its environment. Whether and how an organization recognizes its environment can depend on its information

management systems. How an organization collects, screens and selects information affects the environmental activities that organization will respond to. Pfeffer and Salancik add that "...individuals who attend to the information occupy certain positions within the organization and tend to define the information as a function of their position" (Pfeffer & Salancik, 1978, p. 13). A marketing department will likely view a problem within their scope, just as the finance department might view the same information as a financial problem. Similarly, if the evaluative function is not a designated responsibility within the organization there will be no one to view information within the realm of evaluation. If it resides with the executive director, his or her ability to separate out the evaluation of the program from the evaluation of his or her management and leadership can be a legitimate challenge.

Similarly to the cliché of "what gets counted gets done," we're reminded that "information which is not collected or available is not likely to be used in decision making ..." (Pfeffer & Salancik, 1978, p. 13). This observation has multiple implications for the capacity for program evaluation. On the front end, data collection systems are critical for the ability to conduct evaluations. Without the proper instruments, methods of distribution and collection systems the information necessary to evaluate a program may not be attainable. On the back end, if programs are not evaluated they are most likely to be judged by other criteria such as fiscal efficiency, the leadership's reality construct, and/or history and branding. The information that is gathered by an organization is an indication of what it has deemed important. Typically, only when program evaluation is culturally institutionalized and prioritized is the necessary information collected for the program to be evaluated.

This also raises the issue of *utility*, one of the four main program evaluation standards recognized by the American Evaluation Association. It states “The utility standards are intended to ensure that an evaluation will serve the information needs of intended users” (The Joint Committee on Standards for Educational Evaluation, 1994). The sub standards for utility all point to the fact that evaluations should not be conducted if the information gleaned from them will not be used. An organization that conducts an evaluation may not necessarily use the results to improve the program. It may rather use the results to satisfy a reporting requirement from a funding organization, a government contract or use the information for quality signaling to its constituency.

When examining how individuals in organizations respond to the environment, it is important to remember that individuals have constraints on their behavior. People might have physical, geographical, social constraints or those driven by personal preferences. People within organizations tend to emphasize their responsibilities and allocate time and attention to them. For example, a fund development staff person who ascends to be executive director will likely emphasize the fund development function within his or her managerial sphere of the organization. In extreme cases, the executive director exercises preferences by continuing previous roles or responsibilities while assuming the overall responsibility for managing and leading the organization. Managers’ and leaders’ preferences can be powerful forces they allow to pull them in their desired directions.

The role of management, especially in NHSOs, serves as a representation of the personality of the organization. More so, the symbolism of management portrays it as the focal point of the organization’s successes and failures, activities and outcomes.

Returning to the issue of separation of program evaluation from leadership evaluation, we must wonder if the existing pressure of this symbolism is feared to be compounded through the evaluation of the program. Simply put, will the results of the program evaluation serve as an indictment on the executive director's management and/or leadership effectiveness? This might be especially an intimidating venture for executive directors who have been in their role for many years touting the quality of their program throughout all that time. The revealing of the true value of the program that has been taken on faith for all those years might comprise an "emperor has no clothes" type nightmare for such an executive director. This dilemma is but one component of the overall difficulty all leaders face when they intentionally or unintentionally mix their personal, individual agendas with that of the organization. Managers take risks through this symbolism, as they sometimes are implicated for things they can't control.

The new institutionalism reminds us that the social context or environment of organizations is constantly changing. Competition is an important driver of this changing environment. This helps make yet another case for program evaluation, as NHSO's programs seemingly have to evolve to meet the demand of their consumers within this environment. Program evaluation within NHSOs can aid them in changing their programs to be more effective, more efficient and meet the very needs they were created to serve. It can serve as the framework for developing new and improved methods of program delivery. Some examples might include addressing consumers' transportation needs, personal and work schedules, language and cultural barriers, and personal preferences. This reinforces the notion that programs should evolve to adapt to the environments in which they serve their consumers.

The relationship of the external environment, or “social structure,” to organizations is interactive as organizational variables also impact their environment (Stinchcombe, 1965, p. 142). Some of the characteristics of the social structure that can affect an organization may include institutions, laws, other groups, demographics, and other social parameters that contribute to forming the environment in which an organization exists. This social structure affects organizations in various ways through various channels. For example, there is the rate of the establishment of new organizations. There is a link between when a new organization is formed and the social structure of similar, existing organizations at that point in time. Organizations have stratification systems that determine relationships between stakeholders that not only represent those same class divisions in society but help perpetuate them. Organizations also affect their social structures, typically through the formation of affiliate groups (Stinchcombe, 1965, pp. 143-145).

It is important to understand what some of the motivations are for founding an organization in order to relate them to the context of NHSOs. One reason is to develop a better way of doing things that is not feasible under the current conditions. Founders of NHSOs through entrepreneurial efforts especially feel there is a more effective or efficient way of addressing a particular social issue affecting human beings. Two options are employing a new and innovative method of delivering an existing program or developing an entirely new program to address that social issue. Another reason for founding an organization is that there are indications that it will be supported for its establishment and necessary resources in the future. Founders of NHSOs, while they may feel their new organization is important enough that it will be supported, still face

the uncertainty of the funding world available for nonprofit organizations. We might argue that program evaluation conducted early and regularly might be a way to demonstrate effectiveness and aid in acquiring resources.

Organizations are also founded by individuals or groups who feel they will directly benefit from this new, better way of addressing the social issue. Some NHSOs have been founded by people or groups of people afflicted with the very social problem their new organization addresses. For example, Candace Lightner founded “MADD- Mothers Against Drunk Drivers” 25 years ago after she and her daughters were victims of a drunken driving accident (MADD, 2005). Sometimes the founder is in a position to control the resources necessary to build the organization or in competitive environments they are able to defeat their opposition. On occasion, social entrepreneurs who have excess wealth directly address social issues with their own resources and sustain their organizations for as long as their resources will allow or as long as they are willing to continue contributing their own resources before seeking additional support. More often, though NHSOs get started like the MADD story where people who want to address a social issue they deem important are free to start their own organization and then seek resources from others to develop and sustain it.

Trust is an important factor that determines “whether resources can be moved to innovators...” (Stinchcombe, 1965, p. 147). This point especially applies to new NHSOs attempting to establish such trust. Trust may be based on relationships between founders and other individuals or organizations, but the trust based on the worth, effectiveness and efficiency of this new organization has yet to be earned. Program evaluation, if conducted early and regularly for newly established NHSOs can be a way to demonstrate

efficiency and effectiveness and therefore begin to earn that trust which can help acquire investors and resources.

Establishing new organizations certainly can be a daunting endeavor reflected by their typically high failure rates. Stinchcombe has provided some possible reasons for failure and the challenges new organizations face when he described “the liability of newness” (Stinchcombe, 1965, p. 148). First, roles of stakeholders within new organizations have yet to be established, partly because these stakeholders bring skills attained from outside the organization. Second, efficiency may become delayed when these new roles are forming within a developing structure of rewards and disincentives. Pushing, pulling and overall tension usually results from this evolutionary shakeout that helps build the culture for the new organization. Communication systems are also under construction, so there are many opportunities for misinterpretations, misunderstandings or for information not being channeled correctly from senders to receivers.

There are some functions that can help reduce this liability of newness including “cost accounting, inventory control systems and standard reporting” (Stinchcombe, 1965, p.149). Program evaluation can also help in reducing this liability. The literature from the field of evaluation recommends that program evaluation should be incorporated into program development (Mark, Henry & Julnes, 2000; Mattesich, 2003; Newcomer, Hatry & Wholey, 1994). One common suggestion is to develop a logic model that guides the process of describing the flow of the program, the intended outcomes and the resources necessary to deliver the program. A logic model can also serve as a budget development tool for a program, as once the inputs (resources necessary for the program) are established, their costs can help create the program’s budget. This process at the

inception of the program is logical and most favorable, as it reminds stakeholders why the program was developed, how it operates, what it takes to operate it, the intended results, and how to acquire meaningful, accurate and relevant indicators for these results. When a new program hasn't been tested on an organizational scale in the field, it brings some "liability of newness," so program evaluation can help reduce that liability.

The third reason for this liability is that social actors within and outside of the organization simply do not know each other well. Even for those who come to the new organization as previous cohort groups, the social context in which they must now interact is new to everyone (Stinchcombe, 1965). People that are completely new to everyone are seen as "strangers" and this reminds us that there is an uncertain capacity to trust each other to carry out their responsibilities (Stinchcombe, 1965, p. 149).

The last factor for contributing to this liability is the lack of overall familiarity reducing stability (Stinchcombe, 1965). New people, new products, new services, new structures and new policies all reduce stability within and organization. New organizations don't have the strong ties with customers, suppliers and other stakeholders that older organizations have. These relationships have to be nurtured and developed over time. Customers may at first not demonstrate the same level of exuberance over a new program, service or product that the organization delivering it holds.

The issue of capacity is central to the challenges and overall survival of a new organization. Stinchcombe has reminded us that there are many variables that affect developing capacity within an organization, but elaborates on several he feels have the greatest impact. *Schooling and literacy* simply impact all of the variables, as they are the basis for language and communication, and they drive the ability to develop capacity for

operation and sustainability. *Urbanization* serves as the environment in which new organizations and innovations are drawn to. The social structures of cities “facilitate the formation of new organizations and ease the transfer of customers from old to new suppliers and products.” The *money economy* frees up resources to be used by new organizations and markets to enable customers to patron these new organizations. *Political revolutions* can have a two-sided coin effect in that they may change the interests and priorities that come with a new regime. But, they also can create new resources that may have been unattainable in the old regime. *Organizational experience* exists before the forming of the new organization within individuals, and it can be the main influence for the forming of the new organization. People can draw on their experience to become entrepreneurial and develop new and innovative products or services, and the organizations that produce or provide them. This experience can contribute to organizational capacity through helping to avoid pitfalls, bringing in existing relationships with external organizations and simply contributing a rich knowledge of what has worked well and not so well in the past (Stinchcombe, 1965, pp. 150-152). These issues translate smoothly into factors behind the capacity for program evaluation in NHSOs and subsequent decisions made by the leaders of those organizations.

The rating of organizations impacts the control of resources. Within the stratification systems that we develop for individuals or organizations, we assign ratings that represent the reliability of the social unit. Stinchcombe states that “...it is the social unit, rather than any individual, which has the kind of prestige that can be turned into control over resources on the basis of promised future performance” (1965, p. 172). The

example he uses for context is determining credit ratings for giving loans to individuals, but there are also rankings for organization types such as research for universities, the ability to attract segments of the population for politicians and so forth. These examples are analogous to what might be proposed as a ranking for NHSOs based on promised performance. While some might view this idea as problematic if not dangerous due to the diversity of this sub sector and the sometimes subjectivity of evaluations, a simple standard of whether the program met intended outcomes, and to what degree, might be a defining criteria for ranking NHSOs. This would be in addition to how they utilize their resources which might also be subjective and impacted by the diversity of the organizations within this sub sector.

Etzioni told us in 1964 that “modern society has placed a high moral value on rationality, effectiveness and efficiency” (Etzioni, 1964, p. 1). If this is still true over 40 years later, it begs several questions. One, is program evaluation recognized as an important method that can be used for improving efficiency and effectiveness? Two, if it is then what components and factors drive the capacity for it? Three, if it is then are there enough program evaluations being conducted to represent what Etzioni calls society’s high moral value?

Organizations are “deliberately constructed to seek specific goals” (Etzioni, 1964, p. 3) and in doing so continually evaluate how well they are performing and make adjustments to meet these goals. There are several functions of goals for organizations including establishing a desired future state, parameters for activity and legitimacy. One of the more relevant functions of goals in relation to program evaluation is “serving as

standards for which members of an organization and outsiders can assess the success of the organization – i.e. its effectiveness and efficiency” (Etzioni, 1964, p. 5).

There can be challenges for an organization striving to meet its goals even when those same goals define the organization’s purpose, or mission. Such challenges come in the form of needs the organization acquires as it grows and/or seeks to the demands of its internal and external environments. As organizations adjust to meet these needs, they may put less attention and effort toward meeting its goals, or replace those goals with these needs. This is when “the organizational goal becomes the servant of the organization rather than its master” (Etzioni, 1964, p. 5). The classic tension between an organization’s mission and satisfying its financial needs is one example of this and one applicable to the world of NISOs. This is not to say that both cannot be accomplished, however the push and pull between the two sometimes competing activities is quite common in the uncertain world of funding nonprofit organizations.

Effectiveness is determined by “the degree to which the organization reaches its goals” (Etzioni, 1964, p. 8). He differentiates it with the other “e,” efficiency, which is measured by “the amount of resources used to produce a unit of output” (Etzioni, 1964, p. 8). We’re reminded that measuring each can be problematic. For organizations whose goals are continuous or for whose results being measured are not material, effectiveness is difficult to validate. This can be true for NISOs, especially those whose outcomes are difficult to measure. Typically, the more complex the program, the more difficult it is to measure its outcomes. Efficiency can also be difficult to measure where again the output of services is not material i.e. producing a tangible product. Despite Etzioni’s balancing of the “two e’s” over 40 years ago, there still exists ongoing confusion and

interchangeable use of these terms today in the for-profit, government and nonprofit sectors.

Goal displacement is when an organization substitutes a legitimate goal aligned with its mission with one that is not. Etzioni stated that the most common example is when “an organization reverses the priority between its goals and means” (Etzioni, 1964, p. 10). This can occur in a NISO when the priorities become the number of consumers served or dollars raised (means) instead of how the consumers are impacted by the program (goals or outcomes). An extreme case cited is when an organization and its initial purpose are displaced by some other goals from an interest group within the organization. One example of this might be growth for the simple sake of growth, the inherent belief that all organizational growth in our society is good. Another might be the goal to increase the organization’s budget as a means to acquire prestige or raise people’s salaries. Yet another may be dropping programs that are core to the mission but do not yield the interest group’s desired financial rate of return. Conversely, the organization may develop new programs to attract new funding sources that are not related to their mission.

The founders of an organization can take steps to diffuse the tendency towards goal displacement. One would be to emphasize the direct service toward its goals (outcomes) over the service toward its means (inputs, activities, outputs). The hierarchy and its corresponding power structure can be designed in a way to give power to those involved in that direct service to the goals. In addition, the means activities that directly service these goals can also be prioritized. Physical locations of departments and personnel might be decided based on who directly services the goals. This presents an

interesting possibility for the founders of a NHSO who wish to create a structure and foster a culture for continuous improvement.

While the goal model used for measuring effectiveness and efficiency rests completely on the extent to which the goal is reached, the system model examines the means and process used for attaining the goal. While this approach is more thorough, it is also more costly when used in research or evaluation. An effective program evaluation would not just determine whether the intended outcomes were attained or not; it would also look at the program delivery process as determined in the program theory and corresponding logic model. We can suggest that system models were the precursors to the modern day logic models used today where the program is the system. Etzioni referenced two types of system models, the survival model and the effectiveness model (Etzioni, 1964, p. 19). The effectiveness model would more closely resemble the modern day logic model, as it is designed to display the relationships between the components that would optimize the effectiveness in the attainment of a certain goal (outcome).

March and Simon are cited by Etzioni for their work on organizational decision making. They dispelled the previously held notion that organizations always “search for the optimum rational behavior.” They proposed that organizations rather “seek a satisfying solution rather than the optimum one” (March & Simon in Etzioni, 1964, p. 30). The organization determines the behavior deemed acceptable and will only change that behavior when performance does not measure up to that level. This implies that such motivation is internally driven. I would add that another trigger may be an external demand for improved behavior and the standard set by that external stakeholder.

For NHSOs, that may be a private foundation, community foundation, government agency, United Way, or accrediting body.

Etzioni discusses Max Weber's work that looked at how power was distributed through the bureaucracy of an organization. This work produced the key question of how to control organizations and their participants to maximize effectiveness and efficiency. The challenge is to also minimize the potential undesirable fallout of exercising such control. For an organization to operate effectively, it must set norms and enforce them. There are various ways organizations can use levers such as financial and non-financial rewards and disincentives to enforce such norms.

Exercising such power can have drawbacks, as the subject being controlled is being driven by extrinsic incentives from external forces. This can be analogous to a NHSO that responds to external demands for financial or legal reasons, but does not sincerely have the intrinsic motivation to meet them. Conversely, when an organization internalizes the demands the cooperation will be more sincere and effective. Weber described legitimation as when an organization accepts an external use of power that is in line with its values (Etzioni, 1964, p. 51). This provides a context for when an external organization requires a NHSO to evaluate its program, and the evaluation not only meets the demand but is in line with the NHSO's mission culture and values.

Institutions and formal organization are separate, as the former were aligned with values and commitment, while the latter were seen as having the rational pursuit of goals (DiMaggio & Powell, 1991). Program evaluation proposes an interesting meshing of these two concepts by accomplishing program goals through evaluation and demonstrating the commitment to the organization's values through the evaluation effort

and use of results for improving the program. When organizations exhort their values and commitment, this can be nothing more than rhetoric without the accomplishment of goals demonstrating such commitment to those values. When the leadership of an NHSO creates and fosters a culture within their organization that perpetuates this separation, the values and commitment seemingly are used to substantiate and validate the organization and its program, therefore making it appear its program does not need to be evaluated. In conclusion, values and commitment can be substituted for true accountability when they are not questioned.

Human Services

The term *human services* has become more familiar over the years with the general public, and especially public and nonprofit managers, since the U.S. Department of Health & Human Services evolved out of the disbanded U.S. Department of Health, Education and Welfare in 1980 (Bresnick, 1983). Human services inherently serve people, known as *consumers* (sometimes referred to as clients), as opposed to producing a product (Gardner, 2006; Gibelman, 2003), through “...the act of people helping other people meet their needs in an organized social context” (Cimmino, 2004, p. 6) or similarly “...facilitating clients’ efforts to grow and change while also effectively negotiating the service system in order to meet their needs” (Diambra, 2004, p. 24).

Consumer growth, change, and met needs are the intended results from the interaction of the NHSO, their professional service workers and consumers which is called an *intervention*. To intervene in a consumer’s life can be an attempt to alter either their condition and/or their environment, and is defined generically as to “interfere with the outcome or course especially of a condition or process (as to prevent harm or improve

functioning)” (Merriam-Webster, 2005). Within the more specific context of human services, an intervention is what is delivered to the consumer for the purpose of generating the desired change, or preventing circumstances, for that consumer. This can be in the form of a product such as a transportation voucher, food or a smoke alarm and/or a service such as training, counseling, or psychotherapy. Consumer participation may be involuntary as in the case of a government agency, such as the Department of Family and Children Services, mandating that a parent attend an anger management program or voluntary as in the case of a girl joining the Girl Scouts. A human services program may contain one intervention or several that are delivered to achieve the desired results for its consumers (Brun, 2005). The intervention establishes the basis for social science research and evaluation; bridges theory, research and practice; and is the vehicle for implementing evidence-based practice, all of which will be discussed later.

Human services are primarily provided through organizations due to technology, legislation, funding streams, societal priorities, and other environmental factors contained in markets and hierarchies. NISOs have the common purpose of improving people’s lives (Barker, 1999; Holland, 1995; Cimmino, 2004; Eriksen, 1977; Mehr & Kanwischer, 2004). The human services sub sector of the nonprofit sector is one of the major groups of 501 (c) (3) charitable nonprofits as classified by the National Center for Charitable Statistics (NCCS), and NCCS in January of 2008 reported it had 119,203 organizations filing their tax form 990 to the Internal Revenue Service (IRS) for the most current fiscal year (NCCS, 2008). This amounted to 19.8% of all charitable organizations filing their 990s during this period, the largest representation from any subsector in the United States. NISOs comprise a diverse group of organizations including those in Crime,

Legal Related; Employment, Job Related; Food, Agriculture, and Nutrition; Housing, Shelter; Public Safety; Recreation, Sports, Leisure, Athletics; Youth Development; and Human Services - Multipurpose and Other (NCCS, 2008). Within these main categories are various subtypes classified by NCCS in their National Taxonomy of Exempt Entities (NTEE) (NCCS, 2008). Some examples include Imperial Valley Food Pantry, in California (K31-Food Banks, Food Pantries); Girls Incorporated of Central Alabama (O50—Youth Development Programs); Bosma Enterprises, in Indiana (J30—Vocational Rehabilitation, includes Job Training and Employment for Disabled and Elderly); Texas State Affordable Corporation (L20—Housing Development, Construction, Management); Hillside Resource and Management Corp, Inc., in Massachusetts (I44—Prison Alternatives); SAFEHOME, Inc., in Kansas (P43—Family Violence Shelters and Services); and Minnesota Special Olympics (N72—Special Olympics) (GuideStar, 2008).

These examples are a small representation of the diverse organizational missions, targeted issues, service delivery methods, leadership, structures, staffing, resources and recommended practices among these organizations. Within this great diversity, however there are some commonly cited characteristics for NHSOs. First, their service delivery (practice) is value-based (Brun, 2005; Gardner, 2006; Hull, Jr. & Kirst-Ashman, 2004; Schuerman, 1983; Thompson, 2000). Second, their work has become more professionalized over time and continues in that direction through organizational accreditation; government funding, contractual or regulatory requirements; advanced education; certification; professional association standards and advances in technology (Brun, 2005; DeRobertis & Saldarini, 2004; Diambra, 2004; Gardner, 2006; Leighninger, 2002). Human service organizations have been impacted by a managerialist or more

“business-like” approach to their work, and we find evidence in more common references to functions such as marketing, having well-defined outcomes for consumers, and ‘maintaining focus on their core business’ (Gardner, 2006). Third, they have been greatly impacted by the devolution of government where more services have been contracted out or entirely deferred to NHSOs (Ferris, 1993; Fredericks, Carman & Birkland, 2002; Gann, 2001; Gardner, 2006; Van Slyke, 2003). Lastly, the increasing call for accountability has especially impacted the human services sub sector due to such evolving relationships with government and accreditation organizations which are placing demands for organizational and programmatic efficiency and effectiveness.

The National Organization for Human Service Education (NOHSE) and the Council for Standards in Human Service Education (CSHSE) jointly describe human service workers as “people who hold professional and paraprofessional jobs in such diverse settings as group homes and halfway houses; correctional, mental retardation, and community mental health centers; family, child, and youth service agencies; and programs concerned with alcoholism, drug abuse, family violence, and aging” (NOHSE & CSHSE, 2004, p. 123). Representatives of NHSOs that provide direct service to consumers are sometimes referred to as social workers, direct service workers, or clinicians; however the terms *human service professionals*, *practitioners* or *program staff* will be used interchangeably in this research to encompass the various professions found in these organizations. Program staff, when conducting this intervention may take on various roles including counselor, educator, broker, case manager, mobilizer, mediator, facilitator or advocate among others (Hull, Jr. & Kirst-Ashman, 2004). They are typically assigned to a specific program, in which they use their knowledge and skills to

implement it by providing a specific service designed to address the specific human need (Gibelman, 2003; Schuerman, 1983).

A program has been defined more generically as “a set of resources and activities directed toward one or more common goals...” (Newcomer, Hatry & Wholey, 1994, p. 3). It has been defined more narrowly and appropriately for the context of human services as “...an organized effort to enhance human well-being...” (Chen, 2005, p. 3); or “...an ongoing, planned intervention that seeks to achieve some particular outcome(s), in response to some perceived educational, social or commercial problem” (Fitzpatrick, Sanders & Worthen, 2004, p. 54). Programs involve people, resources, management, and environmental forces (Fitzpatrick, Sanders & Worthen, 2004; Gibelman, 2003) and “are the main vehicles in modern society through which all kinds of formal services are provided to people” (Pawlak & Vinter, 2004, p. 3). Human service programs are developed to address the social needs of human consumers. Pawlak and Vinter describe a human service program as being, “...administered by a private nonprofit or a government organization through designated program personnel who engage in services that include direct interactions with persons receiving the service, within a particular locale, and under certain conditions” (2004, p. 4). Martin and Kettner offer four criteria that describe a human service program and what is intended to accomplish:

1. Addresses an identified social problem
 2. Represents a significant portion of the total activity of an organization
 3. Has goals and objectives (either formally stated or implied), and
 4. Has designated resources, including personnel (because no activity or endeavor can take place without resources)
- (1996, p. 22)

The efficacy of human service programs has been described as comprising quality and cost, however, “The real test of quality in the delivery of social service is what happens to the individual client” (Gurin & Friedman, 1989). Several steps for human service practice are cited in the literature and will now be discussed one at a time. First, the obvious task is to identify and focus on a particular social problem the consumer or consumers, as in the case of a family, are encountering (Schuerman, 1983). Such focus should strive to understand the nature of the problem and articulate it with clarity to other key stakeholders in the service delivery process. If the problem is not clearly understood, the next step of elucidating the objectives of and specifying the procedures for the intervention will not be possible. Objectives need to be clearly stated, so it can be determined whether they have been met through the process of *consensual validation* where more than one practitioner observes the intervention (Schuerman, 1983). This process of identifying the object of change, similar to the program’s theory of change which will be discussed later in the program evaluation section, can be challenging for reasons of feasibility, the complexity of human social problems, and the fact that programs can be fluid and take place over a period of time therefore presenting opportunities for intervention objectives to change.

NHSOs, as a result of increasing competition for resources, increasing pressure for accreditation and the devolution of social services from government to the nonprofit sector, are currently operating in an environment that is increasingly demanding program evaluation. Beyond meeting these external demands, some suggest that NHSOs are obligated to maximize the effectiveness of their programmatic interventions for the consumers and other stakeholders invested in the programs. Bresnick proposes that

fundors and regulators share in this responsibility, as they “...constitute the support network within which these service delivery organizations function” (1983, p. 13).

This sub sector contains great diversity with respect to service delivery types and programs ranging from youth services to personal social services to emergency assistance. Human services have grown over the past several decades due to the commitment of philanthropy and voluntarism (Bresnick, 1983), but predominantly due to the devolution of government (Fredericks, Carman & Birkland, 2002) where services have been deferred entirely or contracted to NHSOs. Human services remain prominent in American society today due to our continual cultural pursuit to improve our lives, and organizations have been developed and expanded to satisfy this desire. Gardner while discussing human services practice describes evaluation as a way of “finding out more about what is happening in practice, how practice is perceived and experienced by the different people involved, and what difference it makes” (2006, p. 233). The choice to focus on human services for this research is based on 1) the core reason for why NHSOs exist which is to improve a person’s life in some way and 2) how their current environment is increasing the importance of program evaluation.

Program Evaluation

There are several definitions of *evaluation* in the existing literature. Mattesich defines it as “A systematic process for an organization to obtain information on its activities, its impacts, and the effectiveness of its work, so that it can improve its activities and describe its accomplishments” (Mattesich, 2003). Mark, Henry and Julnes define evaluation as the “...conduct of systematic inquiry that describes and explains the policies’ and programs’ operations, effects, justifications, and social implications” (Mark,

Henry & Julnes, 2000). Weiss defines it as “the systematic assessment of the operation and/or outcomes of a program or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy” (Weiss, 1998). Fitzpatrick, Sanders and Worthen define it as “...the identification, clarification, and application of defensible criteria to determine an evaluation object’s value (worth or merit) in relation to those criteria” (2004, p. 5). Guba and Lincoln suggest that there is no correct way to specifically define evaluation because such definitions are social constructions and “...it would forever put an end to the augmentation about how evaluation is to proceed and what its purposes are” (1989, p.21).

Mark, Henry and Julnes argue that the ultimate reason for evaluation is for social betterment which they describe as reducing or preventing social problems, improving social conditions and alleviating human suffering (Mark, Henry & Julnes, 2000). Mark, Henry and Julnes suggest that evaluation’s role in the path to social betterment typically takes hold when social programs evolve from the individual determination of a human need, for example as when entrepreneurs found NISOs and develop programs to address a specific human need in society (2000). They argue that without the desire for social betterment, there would be no reason for establishing social programs or for evaluating them (2000, p. 7). They explain that, “By addressing the ends and means of social interventions, evaluation contributes to the attempts to define and realize social goals, to meet human needs, to promote social betterment” (Mark, Henry & Julnes, 2000, p. 21). Mark, Henry and Julnes state that there are four main purposes for evaluation that impact this process and they include assessment of merit and worth, program and organizational improvement, oversight and compliance and knowledge development (2000, p. 13).

Evaluation focuses on the “object of analysis” or “evaluand,” the extent to which results can be generalized, the level of its comprehensiveness regarding data coverage of the effectiveness, efficiency, unintended results and their relationship with the need for what is being evaluated (Tripodi, 1983). This research focuses on a NHSO’s program as the evaluation object. Some examples of NHSO programs include youth development, employment readiness, credit counseling, adoption services, senior citizen housing and others. There are also various definitions for program evaluation including:

- “...the identification, clarification, and application of defensible criteria to determine an object’s value (worth or merit) in relation to those criteria” (Fitzpatrick, Sanders & Worthen, 2004, p. 5)
- “...the application of ...approaches, techniques, and knowledge to systematically assess and improve the planning, implementation, and effectiveness of programs” (Chen, 2005, p. 3)
- “...a systematic process for an organization to obtain information on its activities, its impacts, and the effectiveness of its work, so that it can improve its activities and describe its accomplishments” (Mattlesich, 2003, p. 3)
- “...the systematic collection of information about the activities, characteristics, and outcomes of programs for use by specific people to reduce uncertainties, improve effectiveness, and make decisions with regard to what those programs are doing and affecting” (Patton, 1997, p. 23).

The similarities in these definitions are worth noting. First, program evaluation is a systematic process (drawing on information systems) for collecting information about the program. Second, it involves assessment, judging the worth or value of the program. Third, the ultimate purpose of program evaluation is to gather information used to improve the program. This purpose of improvement is congruent with the premise that evaluation’s primary purpose is to contribute to social betterment. The evaluation of human service programs places a value on that program’s intervention (Sheppard, 2004). As a result, the common goal of social betterment is found in program evaluation’s desire

to improve the program that delivers the intervention which seeks to improve someone's life.

Program evaluation seeks to answer questions related to the effectiveness of a program. These questions found in the literature tend to be grouped into three levels with the caveat that they are not exclusive, they overlap, and they are driven by who is interested in the answers within a political environment.

The *societal level* comprising elected officials, government agencies, large private foundations, the general public, and the media are typically concerned with programs' broader impact on social problems and implications for policy. These stakeholders seek answers to the questions about which services are producing acceptable results and which are not; what is their impact on the social problem; who's lives have been improved by these services; how have funds and other resources been used in these services to address the specific social problem; and how do the benefits compare to the costs (Fitzpatrick, Sanders & Worthen, 2004; Newcomer, Hatry & Wholey, 1994).

The *organizational level* comprises the top administrative stakeholders interested in the results of a program and how they can be used from a more bottom-line perspective. The funder's representative, such as a foundation program officer, United Way staff person, or government agency administrator is interested in the results because they're responsible for reporting them back to their organization. The ED and Board of the NHSO review the results in strategic planning sessions, board meetings, etc. to make organizational level decisions, utilize it for public relations or for acquiring funds. The accrediting organization's representative reviews the results to determine if the NHSO has met the requirements for accreditation. Similarly, the professional association for a

particular human services profession, such as the National Association for Workforce Development Professionals (NAWDP), might be interested in results from their member NHSOs if they are conducting research, revising recommended practices or utilizing information for advocacy efforts.

The *program level* comprises stakeholders such as the program director or manager; program staff and volunteers that deliver the program; and the consumers, especially in the case of empowerment evaluation where they play a primary role in the evaluation process. They are typically concerned with issues specifically centered around the program such as its design, theory, strengths, weaknesses, logistics, context, and process (Bamberger, Rugh & Mabry, 2006; Boulmetis & Dutwin, 2005; Brun, 2005; Chen, 2005; Festin & Philbin, 2005; Fitzpatrick, Sanders & Worthen, 2004; Mark, Henry & Julnes, 2000; Newcomer, Hatry & Wholey, 1994; Patton, 1997; Preskill & Torres, 1999; Russ-Eft & Presskill, 2005). The questions below collectively form the rationale for the program and corresponding evaluation and serve as the foundation for the evaluation plan. They include, but are not limited to:

- Does the program impact consumers by addressing their human needs and resulting in the intended outcomes?
- How satisfied are the consumers with the program?
- Is the program delivered in a manner that respects the consumer's confidentiality, dignity and rights?
- How many consumers are impacted as intended by the program with the given resources?
- Are there any unintended results from delivering the program that are desirable or undesirable?
- What aspects of the program will the consumers indicate helped them the most?
- How effective are the components of the program?
- Under what conditions is the program most effective?
- What are the resources necessary to deliver the program?
- How efficient is the program delivery with regard to utilizing resources?

- What is the program's theory and how well does it match the results?
- How do consumers learn about the program?
- What changes should be made to the program?
- What are the challenges and constraints for delivering the program?
- How do the program's results compare to those of other organizations with similar programs?
- What are the training and technical assistance needs for the stakeholders who deliver the program?
- What can stakeholders learn about the program and their work? (organizational learning)

Some of these questions may be answered through *formative evaluation*, sometimes referred to as *process evaluation*, which is evaluation that provides information related to the design, development or implementation of a new program (Monette, Sullivan & DeJong, 2002; Reamer, 1998; Scheirer, 1994; Schuerman, 1983; Tripodi, 1983). Formative evaluation is concerned with information about a program's operations usually for the purpose of comparing what is being done in the program to the original plan and objectives in order to make modifications. Reamer calls it "*temperature-taking evaluation*" because it contributes to the monitoring of the "health" of the program's plan of action (1998, p. 277). The purpose of the methods used in formative evaluation is to serve as a pilot test run with what would be representative of program participants to see how they interpret the program, how they are affected by it, and if it is conducive to their cultures (Scheirer, 1994). Examples of types of data sought in this process may be information about the program's intake process, the duration of the consumer's participation in the program, tools and methods used to communicate with consumers, the effort to provide accessibility to underserved or special populations, or time allotments for key staff functions or stages in the program.

Needs assessments are examples of formative evaluation quite common in human services when organizations attempt to acquire a knowledge base of consumer needs that goes beyond an existing body of knowledge based on the literature, past research and evaluations. Some methods may include one-on-one interviews, focus groups, surveying a sample of the population to be served, testing equipment, conducting a communications analysis, observations, or pre-testing an intervention. They may be combined for a mixed-method approach which has advantages for acquiring different types of information, incorporating different perspectives, addressing the pitfalls of using only one method, and utilizing the methods most appropriate for a particular participant population, environment, or context.

Some questions may be answered through *summative evaluation* which is more structured than formative evaluation, and is concerned with program outcomes and related issues of generalizability, replicability and inference (Boulmetis & Dutwin, 2005; Reamer, 1998; Schuerman, 1983; Tripodi, 1983). It also helps satisfy the stakeholders who are interested in the more “bottom line” answer to the question of whether the program’s results significantly affected the consumers as intended. Summative evaluation helps provide internal stakeholders such as the ED, program staff and board of directors with information about how well the program is meeting its objectives. This can be done on a broader scale for comparison if the NHSO delivers the program across multiple sites. It also helps provide information to external stakeholders that have a direct interest in the program such as those who fund it, contract for it or determine accreditation for the organization delivering it.

The main purpose of summative evaluation is to provide a value judgment of the program's merit and worth with regard to its intended or unintended consequences and other characteristics such as the protection of participants' rights (Mark, Henry & Julnes, 2000). Mark, Henry and Julnes, in keeping with their overarching theme and primary purpose for evaluation, social betterment, recommend that values such as fairness, equality, justice, and fiscal responsibility be incorporated into this process to go beyond just the specific intended results of a program (2000, p. 43). This perspective includes societal judgments and is consistent with the view that program evaluation goes beyond just searching for the attainment of the program's goals because the goals themselves also need to be evaluated (Palumbo, 1987; Scriven, 1993).

Scriven discusses some of the misperceptions about program evaluation by providing several theses in an attempt to provide a more well-rounded perspective for what it is and should be. He posits that describing program evaluation as determination of goal attainment undercuts its purpose by leaving out evaluative statements about the goals or merits of the program (Scriven, 1993). In other words a NHSO's programs can meet or not their goals, but what if those goals were not realistic, soft, not established based on past experiences or designed to solve effectively the wrong problems. The goals can also be based on a well constructed and conducted needs assessment that leaves a substantial time gap between it and the implementation of the program. Goal attainment may also not address ethical issues, such as those relating to the consumers being used as human subjects. It also will not likely address the side effects of the program such as unintended outcomes, the relationship between the size of the program and its effectiveness among other issues. Scriven tells us that program evaluation "must

go beyond acceptance of the goals into evaluation of them, it must consider costs and comparisons, and it sometimes benefits from ignoring goals entirely, except in an appendix to an evaluation report” (1993, p. 19).

Program theory serves as a framework for understanding how a program operates and why it should result in its intended outcomes (Bamberger, Rugh & Mabry, 2006; Fitzpatrick, Sanders, & Worthen, 2004; Mattesich, 2003; Rogers, Petrosino, Huebner & Hacs, 2000; Thompson, 2000). Three major functions of theory that establish this framework include helping to explain phenomena in terms of what will happen and why it will happen; guiding research and practice by directing stakeholders towards finding answers about the program they are studying; and enabling stakeholders to integrate and analyze results in various environments and contexts (Monette, Sullivan & DeJong, 2002). Program theory plays an important role in human services and the evaluation of programs by helping those invested in human services understand individuals and their issues in broader social and political contexts and structures through critical thinking about people and their social problems (Mattesich, 2003; Thompson, 2000).

Program theory is an important aspect of the evaluation of human service programs and human service practice in general. Lipsey states that theory is necessary for knowledge because it helps us understand why events happen and upholds new insight into solving a particular social problem (1993). He elaborates, “Theory-oriented research holds out the promise of increasing knowledge in ways that build the practical science of social intervention while informing policy and practice throughout the helping disciplines” (1993, p. 35). Three types of theory are relevant for human services practice

and the evaluation of programs. *Problem theory* addresses the causes of the specific social problem, such as what causes gambling addiction or homelessness. *Outcome theory* is concerned with identifying a desired social outcome in lieu of the current state of a particular social problem and what social actors can transition the problem to the desired outcome (Weiss, 2000). Once the problem, the desired outcome and the social actors are established, a *treatment theory* or *theory of intervention* can be established to determine the logic for action, how the social actors will address the problem to result in the desired outcome. *Treatment theory* addresses what can be done to change the condition of the problem, such as what can be done to help a person to eliminate their gambling addiction or get them off the streets and into a shelter, transitional housing, or permanent housing. It is concerned with using inputs to transform the undesirable current state of a problem into a more desirable output (Lipsey, 1993). The required elements of treatment or intervention theory “are the agent (who should intervene), the target (whose actions are to be changed in some way), the mechanism (how to intervene), and the time and place (when and where a concrete social intervention takes place)” (Weiss, 2000, p. 86).

In the context of evaluating nonprofit human service programs, problem theory represents the current body of knowledge about a specific human, social problem based on prior research, evaluations, experiments, the literature and the knowledge captured from practice. Outcome theory represents the desired state or condition of the human consumer who is receiving the intervention. Treatment or intervention theory takes the knowledge from problem theory and combines it with outcome theory to determine what the most effective intervention is for addressing the problem to attain the desired outcome. Together, they comprise the program theory that is the foundation for how a

human services program operates and intends to impact the problem and result in desired outcomes.

Program theory has operational or logistical implications for the evaluation of human service programs. It draws from a business management used for many years for setting goals and measuring progress towards attaining them. First, it drives the methodology for a particular program evaluation. It helps an evaluator determine what is the most effective and efficient methodology to use to evaluate the program that incorporates variables, constructs, the desired outcomes, the treatment and the relationship between the treatment and those outcomes (Lipsey, 1993). Program theory also applies to individual stakeholders. For example, how human service professionals see their role in the intervention process impacts the outcomes of that intervention (Thompson, 2000).

Program theory also helps reduce the potential for discriminatory practice, which could occur subconsciously when human service workers rely on their intuition or common sense, and would be counter to the values and ethical standards for human services professions (Thompson, 2000). It is important to note however that this may not always be the case, as theory itself can be discriminatory or contain bias resulting in the mistreatment of consumers as human subjects. Another way of realizing the importance of program theory is that consumer needs, corresponding consumer needs and programs are not likely to remain static because the environment is constantly changing. Program theory therefore is also not static and serves the purpose of being renewed from the latest research, evaluation results and innovations in service delivery. The importance of

renewing and refining theory and helping it to evolve through these methods lies in reducing the potential for such circumstances to occur as well as repeating past mistakes.

Logic models are visual depictions, usually in a flow-chart form, of a program's theory of change and how activities will result in intended outcomes. They are an important tool that helps NHSOs convey their expectations for their program by connecting interventions to desired results (Brun, 2005; Eliason, 2007; Monroe, Fleming, Bowman, Zimmer, Marcinkowski, Washburn, et al., 2005). There may be cases where human services programs have been delivered for many years without ever there being an effort to discuss and formalize their theory of change. Logic models also can help NHSO stakeholders understand the evaluation language and such terms as inputs, outputs, outcomes, continuous improvement and others (Monroe, Fleming, Bowman, Zimmer, Marcinkowski, Washburn, et al., 2005; United Way of America, 1996). Logic models help NHSOs develop their outcome measurement framework and subsequent process evaluation (Eliason, 2007).

Scholars and researchers often recommend that the logic model development process begins with the research questions and focuses on indicators for the intended outcomes (Eliason, 2007). Logic models ultimately help an NHSO answer the questions for how program outcomes will be defined and measured, and the what, when and how for data collection (Eliason, 2007). In a more practical sense they help NHSOs answer the following questions for the outcomes they have determined they want to measure in their program evaluation process:

- Are they relevant to the organization's mission and objectives of the program?
- Do they represent realistically what the program should be accountable for?

- Do they represent changes or benefits to consumers that are meaningful and relevant to the social problem being addressed?
 - Is there a theoretical but clearly stated connection between the program's activities and intended outcomes?
- (Adapted from United Way of America, 1996, p.57)

Scholars, researchers and practitioners indicate that these contributions logic models can provide establish them as valuable tools for ECB (Monroe, Fleming, Bowman, Zimmer, Marcinkowski, Washburn, et al., 2005; United Way of America, 1996). Others suggest that the logic models contribute to mainstreaming evaluation by enabling stakeholders to verbalize their program's expectations (Barnette & Wallis, 2003). They also reiterate that it's not necessarily the formally established model itself that is so valuable. It is the process in which they are used that involves stakeholders; stimulates their interest, enthusiasm and learning for the program; and gives new or renewed meaning to their work as they realize their roles in consumer outcomes. Logic models are consistently referenced in the majority of ECB approaches which will be discussed in the next section (see *Approaches*).

There are some current trends around program evaluation that NHSOs should be aware of. Evaluators were once perceived as being objective or neutral for what they were evaluating to retain the integrity for the profession and for the evaluation process. However, evaluators like evaluations are not free from values, as they, too are engaged in social construction even with the use of scientific methods (Fitzpatrick, Sanders & Worthen, 2004). The science and art of making a value judgment of something clearly cannot be value free. Also, the independence of the evaluator can be more of a perception that is largely driven by the stakeholders of the organizations contracting with them who are in a position to strive to create an environment in which evaluators are

encouraged and allowed to be “independent” (Reingold, 2006). If such an environment is not created, the personal ethics and adherence to the standards and guiding principles of the evaluator may not be conducive for an assignment’s acceptance or continuance.

Another trend that continues is the increasing awareness and knowledge of evaluation tools and methods. This has resulted from publications like the United Way’s *Measuring Program Outcomes: A Practical Approach* in 1996, the W.K. Kellogg Foundation’s *Evaluation Handbook* published in 1998, and more recently Bamberger, Rugh and Mabry’s *Real World Evaluation* and Festin and Philbin’s *Level Best: How Small and Grassroots Nonprofits can Tackle Evaluation and Talk Results*. These publications contribute to an increasing trend in the evaluation profession of demystifying evaluation so an organization’s stakeholders can break their misperception of it being a foreign, scientific concept, learn enough about it to effectively work with evaluators, and overall work towards building internal evaluation capacity (Duigan, 2003). As evaluation capacity building continues to gain momentum as a legitimate concept and helpful process for NHSOs interested in evaluating their programs, the task of demystifying evaluation will continue to expand in presence and importance. The growth and evolution of the evaluation profession has also contributed to the increasing awareness and knowledge of evaluation tools and methods, and this progression and its byproducts will be discussed later in the section about the role of the evaluator.

The growth of evaluation on an international level is one of twelve emerging trends or issues Fitzpatrick, Sanders and Worthen believe will influence program evaluation that also includes “Increasing priority and legitimacy of internal evaluation” (2004, p. 44). This can be observed from the increasing number of evaluation staff

positions at most large NHSOs, especially at national headquarters offices such as the Girl Scouts USA or Boys & Girls Clubs of America where entire evaluation teams or departments exist.

Another important trend in program evaluation for NHSOs is the increasing incorporation of qualitative tools and methods. The past argument between the quantitative and qualitative camps has dwindled, as the evaluation literature more often references the importance of using complimentary, mixed methods. This is reassuring to NHSO stakeholders, especially those in the human services professions such as social work, counseling and psychology that utilized qualitative methods such as interviews, observations, focus groups, etc. for many years. The argument for mixed methods we hear more prominently today than in the past reminds us that the survey results in the form of numbers and percentages or the consumer one-on-one interviews alone would not be sufficient to holistically evaluate a program, especially a human services program dealing with complex social problems.

Evaluation is increasingly having a place in academic programs in higher education across the United States. This is contributing to the increasing professionalization of evaluation. The American Evaluation Association in 2008 reported that there were 48 universities offering "...graduate programs or certificate programs either directly in evaluation or with available concentrations in evaluation" (AEA, 2008). Western Michigan offered the first interdisciplinary PhD program in evaluation in 2005. There has also been an increasing presence of graduate courses in evaluation in nonprofit studies, psychology, and social work programs. These academic programs can contribute to the future pipeline of evaluators and human service or

nonprofit professionals having a working knowledge of evaluation so they may participate in the program evaluation process.

Brun states that “Evaluation is social work...” because it is a process that improves social services, represents accountability, and guides practice (2005, pp. 2-7).

Some commonalities between human services practice and program evaluation have been referenced throughout this section that support this claim and are worth noting:

- Being systematic processes
- Requiring resources
- Involving people
- Residing in a socio-political context
- Making a value judgment
- Having the goal of social betterment
- Incorporating needs assessments
- Using qualitative methods
- Incorporating theory
- Focusing primarily on the intervention
- Using results to guide planning and practice
- Building knowledge

Human services practice and program evaluation both are driven by relationships and seek answers to many common questions. The concept of *practice wisdom* represents the knowledge gained by practitioners within a given profession. Practice wisdom is the basis for evidence-based, research-based and reflective practice all which will be discussed further in the *Program Staff* section. Mattesich claims that evaluation is an important complimentary partner for practice wisdom because it helps provide a broader picture of the profession and help reduce the bias driven by that profession’s limited view of itself (2003, p. 9).

Another helpful exercise for examining their similarities is in the comparison of their professional standards and ethical codes. Common principles and standards are shown below in Table 1.

Table 1 – Common Principles and Standards among Professions

Principles/Standards	Human Services				Evaluation
	Social work	Workforce development	Psychology	Counseling	
Informed consent of consumers	X	X	X	X	X
Competence for profession	X	X	X	X	X
Cultural competence	X	X	X	X	X
Privacy/Confidentiality	X	X	X	X	X
Integrity for profession	X	X	X	X	X
Conflict of interest	X	X	X	X	X
Respect, worth, and dignity of consumers	X	X	X	X	X
Social equity/justice	X	X	X	X	X

Sources: Social Work National Association of Social Workers (NASW)
Workforce Development National Association of Workforce Development Professionals (NAWDP)
Psychology American Psychological Association (APA)
Counseling American Counseling Association (ACA)
Evaluation American Evaluation Association (AEA)

The four human service professions in Table 1 are not exhaustive but serve the purpose of representing some of the more prominent ones in the subsector. The eight principles or standards they share in common are also not exhaustive of all in each profession but were chosen to serve the purpose of accentuating the ones they all have in common. We see commonalities concerning consumers, standards for the professions, ethical issues, and ensuring equity and justice for stakeholders. These commonalities illustrate how these professions have evolved, but more importantly for this study, help us understand that evaluation has begun to become an integral part of human service professions.

This relationship can also be seen more explicitly where evaluation is referenced as a principle or standard in the codes of these human service professions. This will be examined in more detail in the *Program Staff* section which discusses their role in the

program evaluation and evaluation capacity building processes. The purpose of these examples is to further our recognition and understanding of the many common and complimentary characteristics of human services practice and program evaluation. They are relevant and important for their potential to serve as a powerful influence on the building of evaluation capacity in a NHSO.

Evaluation Capacity

Constructs

A NHSO's decision for whether to evaluate its programs is likely driven by the following factors:

1. Desires of evaluation sponsors
 2. Focuses of evaluation
 3. Clarity and specificity of programs and practices
 4. Evaluation objectives
 5. Available resources, time and expertise
 6. Potential for using the results of evaluative research
- (Tripodi, 1983, p. 12)

Tripodi's list is one example of issues related to *evaluation capacity* that have been discussed for decades usually for describing aspects of readiness for evaluation.

However evaluation capacity as a specific term, concept and construct is relatively new having permeated the literature as recent as the late 1990's. It is intended to frame what is necessary for an organization to evaluate something and in the context for this research, the ability for a nonprofit human services program to be evaluated. Evaluation capacity is defined as "...human capital (skills, knowledge, experience, etc.) and financial/material resources..." that are necessary for the practice of evaluation (Boyle & Lemaire, 1999, p. 5).

There are several definitions of *evaluation capacity building* (ECB) found in the literature (Bamberger, Rugh & Mabry, 2006; Preskill & Catsambas, 2006; Stockdill, Baizerman & Compton, 2002). Compton and Baizerman attribute this to the fact that ECB has gained momentum since it was AEA's annual conference theme in 2000 and has grown conceptually faster than the actual practice (2007). They describe this as "a 'natural' stage in the institutionalization of both knowledge about and knowledge 'how to' in a practice-oriented profession" (Compton & Baizerman, 2007, p. 118).

The ground-breaking, conceptual definition that Stockdill, Baizerman, and Compton offer helps establish the foundation for the comprehension of ECB:

"ECB is a context-dependent, intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which quality program evaluation and its appropriate uses are ordinary and ongoing practices within and/or between one or more organizations/programs/sites" (2002, p. 8).

Some of these components will be addressed individually, as according to Stockdill, Baizerman, and Compton's breakdown of their definition, and will be complimented with ideas and concepts from other researchers and scholars.

Each environment, organization, and program contains unique issues and factors that can impact ECB rendering the process *context-dependent*. Volkov and King remind ECB practitioners in their *Checklist for Building Organizational Evaluation Capacity* to "Be aware of the internal and external organizational context, power hierarchies, administrative culture, and decision-making processes" (2007). Patton's context focus asks the question, "What is the environment within which the program operates politically, socially, economically, culturally, and scientifically?" (1997, p. 192).

Stufflebeam in his *Institutionalizing Evaluation Checklist* breaks out the contextual factors in detail:

“Since evaluation systems are context dependent, take into account constituents’ needs, wants, and expectations plus other variables such as pertinent societal values, customs, and mores; relevant laws and statutes; economic dynamics; political forces; media interests; pertinent substantive criteria; organizational mission, goals, and priorities; organizational governance, management, protocols, and operating routines; and the organization’s history and current challenges” (2002).

The *intentional action system* part of the Stockdill, Baizerman, and Compton definition serves to describe ECB as purposeful, collaborative, and regularly occurring. This description is analogous to Patton’s utilization-focused evaluation concept which includes “intended use by intended users” (1997, p. 23). Bamberger, Rugh and Mabry describe the intentions of ECB as “an enhancement of the quality and utility of the ongoing evaluation and as an investment in strengthening the use of findings” (2006, p. 164). One example of Volkov and King’s overarching intentional theme for ECB is to “Purposefully create structures—mechanisms within the organization—that enable the development of evaluation capacity” (2007).

ECB is *guided* because it is facilitated to be established and sustained. Stockdill, Baizerman, and Compton point out that ECB practitioners who facilitate the process cannot accomplish this alone and require the efforts of other stakeholders. We see examples of stakeholder involvement in several descriptions of the ECB process. Preskill and Catsambas reference “increasing members’ evaluation competencies,” (2006, p.123); while Bamberger, Rugh and Mabry talk about “strengthening the motivation and capacity of managers, planners, policymakers, legislators, funding agencies, and public opinion to commission, assess, and/or use the findings of evaluations” (2006, p. 164). Volkov and

King support Stockdill, Baizerman, and Compton point's a bit more explicitly in several recommendations:

“Make sure that key leaders of the organization support and share responsibility for ECB. Locate existing and enlist new evaluation champion(s) in the organization. Provide opportunities for sufficient input in decision making, ensuring that people in the organization are able to use data to make decisions” (2007).

Stufflebeam suggests that ECB practitioners “Promote and support stakeholders’ buy-in, participation, and support from all levels, e.g., by engaging representative panels to review evaluation plans and reports and working to assure that top management and governance are knowledgeable, supportive, and involved in the evaluation effort” (2002).

Process is meant to describe the overall effort to have ECB evolve to be systematically intentional and sustainable through efforts such as establishing an evaluation unit in the organization, marshalling and developing the necessary things for the capacity (ability) to evaluate programs, and promoting and ensuring use of the evaluation results. Stufflebeam, and Volkov and King, discuss resources, structures, policies, procedures and communication within the context of sustainable ECB efforts to support ongoing evaluation.

Stockdill, Baizerman, and Compton include *practices* to represent intentional and reflective every day activities that contribute to ECB such as convening stakeholder meetings and providing training. Volkov and King in their ECB checklist offer recommendations that compliment this idea such as “Organize opportunities for socializing around evaluation activities during the workday (for example, working on a survey collaboratively or discussing evaluation findings at brown bag lunches)” and

“Establish clear expectations for people’s evaluation roles and provide sufficient time during the work day for evaluation activities.” (2007).

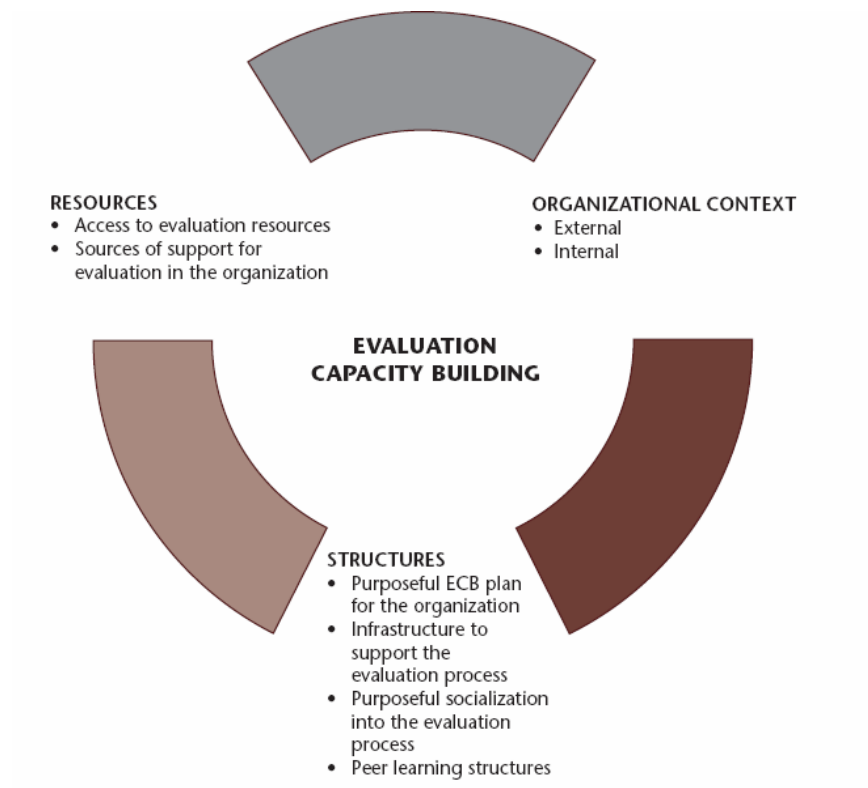
The phrase *ordinary practices* means that evaluation becomes regular, routine, a basic assumption that has been mainstreamed into the culture and operations of the organization. Sanders defines mainstreaming evaluation as “...the process of making evaluation an integral part of an organization’s everyday operations” and “...part of the organization’s work ethic, its culture, and job responsibilities at all levels” (Sanders, 2003, p. 3). Duigan suggests that mainstreaming evaluation is bonded to, if not synonymous with, with ECB. His rationale is based on the notion that in order for the organization’s stakeholders at all levels to be more evaluative in their work, they “must have appropriate evaluation skills, systems, structures, and resources” (Duigan, 2003, p. 12). Williams and Hawkes tell us we can determine evaluation is mainstreamed in an organization “when it is on every agenda, when evaluation involves roles for those not traditionally affiliated with evaluation activities, when buyers ask for evaluation data on all their purchases, when the CEO distributes a list of organizational values that includes continuous evaluation, when orientation for new employees includes their role in evaluating services and products, and when evaluation advocacy is a criterion for orienting new staff” (2003, p. 64). This may seem unrealistic given the inherent challenges NHSOs already face, and mainstreaming may appear to be more of an ideal state where as ECB leans toward ensuring the capability for long-term commitment to evaluation.

The purpose of Stufflebeam’s entire checklist of 18 recommendations is to institutionalize evaluation in the organization as the checklist’s title indicates (2002).

The purpose of Volkov and King’s checklist is explicitly congruent with Stockdill, Baizerman, and Compton’s characterization by referencing “incorporating evaluation routinely into the life of an organization,” “long-term capacity” and “to conduct and use program evaluations in everyday activities” (2007). Preskill and Catsambas in their definition for ECB mention “sustaining evaluation and evaluative thinking as a way of life in the organization” (2006, p. 123).

Baizerman, Compton and Stockdill also provide a much more practical working definition of ECB, “*the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine*” (Baizerman, Compton & Stockdill, 2002, p. 1). Volkov and King offer three main elements for ECB – *resources, structures and organizational context* contained in their ECB checklist and depicted below in Figure 3.

Figure 3 – Elements for Evaluation Capacity Building (Volkov & King, 2005)



These three elements will each be discussed individually and together serve as a guide to further break down their specific components and strategies to provide a more expansive and detailed account for ECB.

Baizerman, Compton and Stockdill state that “The promise for ECB is limited by the lack of necessary financial and human resources” and that resources go beyond the knowledge and skills for evaluation to include what is necessary to facilitate ECB (2002, p. 21). Examples of the necessary *resources* to evaluate a program include people, facilities, funds, equipment, software, and time (Mesch & McClelland, 2006; Stufflebeam, 2002; United Way of America, 1996; Volkov & King, 2005).

Organizations select and effectively utilize internal or external evaluation personnel and their expertise (Sonnichsen, 1999; Stufflebeam, 2002; Volkov & King, 2005; Werther, Jr. & Berman, 2001). There essentially are four scenarios for how personnel can be utilized with factors that drive choice decisions as well as pros and cons for each.

First, an organization can decide to use existing staff provided they have the knowledge, background, skills and time for evaluation designs, methods and collecting, analyzing and reporting data (Mattesich, 2003). There are other factors behind evaluation capacity which will be discussed in upcoming sections that may lead us to believe that these alone will not suffice.

Second, an organization can decide to create a staff position dedicated to the evaluation function and hire a person as an internal evaluator on staff. Senior management support for the evaluation function and available funds for the new position are likely to be the driving forces behind this choice.

A third option is to contract with an external evaluation consultant to handle the evaluation function. A consultant can be hired for various functions at different stages in the process such as designing the evaluation, serving as an occasional advisor or to handle all aspects of the evaluation work (Mattesich, 2003). It is important to note that even if an external consultant is hired to perform all parts of the evaluation process, he or she still would have to communicate with and rely on internal stakeholders, most likely program staff, for important information about the program.

This leads us to the fourth option of purposefully utilizing internal staff with skills and interest in evaluation with external evaluators in tandem to evaluate the program. On the surface this may appear to be the ideal option for balancing internal program, operational and cultural knowledge with the expertise and fresh perspective of an external evaluation consultant.

These choices may cause anxiety for a NHSO because there is no *Consumer Reports* for evaluators. Stevahn, King, Ghare, and Minnema recently developed a set of competencies for evaluators that NHSOs can use as a guide for what to look for (see *Evaluators* section in Stakeholder Roles). Additional resources such as the program evaluation standards and the guiding principles for evaluators can help a NHSO better understand what to expect from an evaluator and the evaluation profession. Important internally driven factors include how the evaluator meets the organization's current needs, how well they know the social problem and service delivery, and how familiar they are with the target population the organization serves.

The choice for utilizing an external evaluator may be driven by a NHSO's intention to appear or be objective, satisfy external requirements of funders, government

contractors or accreditation organizations, or reduce costs by saving on employee benefits. The environmental context for NHSOs may render these decisions to be based solely on cost, especially when external stakeholders categorize evaluation as administrative expenses rather than expenses for the program. While this tactic may appear at first to satisfy the feasibility standard for program evaluation, it may prove counterproductive to the ECB process. Quality and sustainability may be sacrificed, a signal is sent to the stakeholders that evaluation is not an organizational priority, and the organization will likely evaluate its programs on an ad hoc basis.

The types of external evaluators that can be contracted with include individual independent consultants, nonprofit or for-profit consulting firms, or faculty or staff from a university or academic research center (Mattesich, 2003). Issues to consider when making this choice include the evaluator's ability to customize their approach to the organization and its program instead of a boiler plate approach, ask critical questions of the organization to better understand its environment and context and make an effort to understand who the users of the evaluation are as well as organizational culture issues. NHSOs may turn to the American Evaluation Association's "Find an Evaluator" section of their web site where they can search for evaluators by name, state or area of expertise (AEA, 2008). They can also issue requests for proposals (RFPs) and post them on list serves evaluators are likely to use to look for work such as Idealist, Opportunity Knocks, or EvalTalk. The W.K. Kellogg Foundation offers tips in their *Selecting an Evaluator* section of their evaluation handbook that cover types of evaluators, the role of the evaluator in program evaluation, how to find an evaluator, evaluator qualifications and includes a checklist for selecting an evaluator (W.K. Kellogg Foundation, 1998).

When a NHSO contracts with an external evaluator, they do not absolve their staff of involvement in the evaluation process. To the contrary staff time will have to be dedicated to and allocated for the evaluation process. Mattesich offers some helpful tips on how an organization can be ready to effectively utilize an external consultant:

- Have a program theory.
 - Intend to use the results of the evaluation.
 - Make your expectations as clear as possible.
 - Develop a good advisory committee.
 - Consider every step to be a collaborative process.
 - Focus on the information needs of the users of the project's results.
 - Budget enough time (for the design and the work itself).
 - Budget enough money.
 - Develop clear and reasonable standards for communication and progress reports.
 - Realize there will be some ambiguity.
- (2003, pp. 68-69)

Funds are an important resource that ultimately can determine whether programs get evaluated. Government and private foundation grants occasionally have line item expenses to cover evaluations, but there is no clear, overall sense of whether these evaluations are being funded at the necessary levels determined by the nature of the programs being evaluated and what kind of information is desired. Scholars and practitioners have observed that evaluation project budgets typically do not have sufficient funds (Bamberger, Rugh & Mabry, 2006). There are several potential scenarios that can demonstrate how the cost of program evaluation combined with the organization's culture for it are driving forces behind decisions whether to engage in it. A NHSO that treats program evaluation as an expensive luxury only for organizations that are well endowed with resources or for their organization only when an external stakeholder will pay for it, the organization is not likely to prioritize it, plan for it, budget for it and work towards long-term evaluation capacity. If an organization has the funds to

pay for program evaluation but the culture doesn't value or support it, it may not occur. Lastly, an organization may not have the funds to pay for program evaluation but the culture supports it enough for them to plan and budget for it.

Some of the costs for program evaluation include those related to internal and external personnel, data collection and analysis, supplies and equipment, travel, and general overhead (Preskill & Catsambas, 2006; Wholey, Hatry & Newcomer, 1994). There are some strategies NHSOs can use to reduce the costs for program evaluation.

Bamberger, Rugh and Mabry suggest the following:

1. Simplify the evaluation design
 2. Clarify client information needs, seeking ways to cut out the collection of nonessential information
 3. Look for reliable secondary data
 4. Reduce the sample size
 5. Use more economic data collection methods
- (2006, p. 51)

The overarching theme of their recommendations is to sharply focus on and narrow down the needs of the organization running the program and prioritize the evaluation. It is important to note that these strategies can be threats to validity and reliability. An organization may have to compromise at some level between reducing expenses and maintaining data integrity, the type of compromise found quite often in research (Andresen, Machuga, Van Booven, Egel, Chibnall & Tait, 2008; Keeter, Kennedy, Dimock, Best & Craighill, 2006; Newcomer, Hatry & Wholey, 1994; Rubin, 2006).

Organizations can also generate revenue to support and assure long-term, dedicated funding for evaluation-related activities (Volkov & King, 2005). The financial status of NHSOs is often affected by government funding cuts, private funder interests changing, grants ending, reductions in corporate giving, increases in service delivery and

other operational costs, as well as an overall downturn in the economy. This can impact a NHSO's evaluation capacity through forced budget cuts that reduce or eliminate available financial resources; staff turnover; reduction in time allocated for the evaluation process; and overall organizational instability (Tripodi, 1983; Weiss, 1998). Capacity building has been a popular phrase in the U.S. nonprofit sector for several decades. NHSOs strive to build capacity to acquire more buildings, hire more staff, and serve more consumers, so they are now faced with the decision for whether they will also build capacity to evaluate their programs.

Time is an important resource for the program evaluation and ECB processes (Fitzpatrick, Sanders & Worthen, 2004) and serves as a major criterion for setting an evaluation's agenda (Newcomer, Hatry & Wholey, 1994). While one of the feasibility standards reminds us that "The evaluation procedures should be practical, to keep disruption to a minimum while needed information is obtained" (Joint Committee on Standards for Educational Evaluation, 1994), time is a resource that must be properly planned for, estimated, and allocated. Its role in ECB is obvious, as if not enough time is allowed even the most qualified evaluator and well intended, knowledgeable stakeholders will not be able to successfully achieve the goals for the evaluation and produce reliable and valid information that can be used.

Bamberger, Rugh and Mabry suggest reducing the amount of effort in stakeholder time or lessening the duration of the evaluation as two broader, distinct strategies for dealing with time constraints (2006, pp. 69-71). The choice for either of these strategies or both depends on the requirements for the evaluation, as a program cycle might dictate

the evaluation must last nine months, however they offer four specific strategies that can be used to save program staff or evaluator time within that ninth month period.

One is to decrease the time burden for external evaluation consultants through maximizing the effectiveness of their time by ensuring they are not using it for tasks that can easily be completed by staff or volunteers. This can present opportunity costs for staff as well as real costs for their time; however the cost of a consultant's time is likely to be higher. Also, by burdening the consultant with these tasks, an organization creates opportunity costs for the consultant with regard to what he or she could be doing for the evaluation that is more geared to their expertise. An organization's leadership can play an import role for reducing time by encouraging and approving the time internal stakeholders, especially program staff, spend on the evaluation process. If the organization's culture supports program evaluation, it is more likely they will allow sufficient, dedicated staff time and less likely it will justifiably view it as an opportunity cost for what they normally be doing, most prominently serving the program's consumers.

The evaluator can hire more people or subcontract some of the work, but this may be challenging to stay within the budget and to maintain consistent quality for the evaluation. A third strategy is to look for opportunities where outcome indicators can be included along with activities and outputs in the project monitoring process. Some examples include consumer changes in attitudes, gained knowledge or consumer access to services, which can also be an outcome in some instances as well as an activity. Finally, Bamberger, Rugh and Mabry suggest maximizing the available technology that meets the project's data collection needs. Some examples include web-based surveys

that dump data directly into a database, recording interviews digitally to produce audio computer files or using voice recognition software to cut down on transcription time, and inputting data directly in real time in the field for reducing post collection data entry (Bamberger, Rugh & Mabry, 2006, pp.81-84). All of these strategies come with cost-benefit analyses as they may put the project over budget, compromise quality or have the overall cost to employ them exceed the time savings and corresponding cost savings.

Structures are referred to here as the physical and mental processes within the NHSO that contribute to ECB. The overall ECB plan is in itself a structure that is used to monitor the success of the process. It also guides and determines other structures the organization can establish for the ECB process, some of which will now be discussed. The establishment of an evaluation oversight group, sometimes called an evaluation work group, team, or advisory committee that can be the catalyst for stakeholder involvement, effectively utilize available human resources, and establish and move forward evaluation processes (King, 2005; Patton, 1997; Volkov & King, 2007; Wholey, 1994). This group should include a mix of program staff, board members, the evaluator and consumers if possible, but could also include consumers, program sponsors, collaborative partners, direct service volunteers, policymakers, local evaluation experts from colleges or universities, or other stakeholders invested in the organization's program. Stuffelbeam reminds us that these stakeholders should possess skills and expertise necessary for the ECB process such as "field work, group process, interviewing, measurement, statistics, surveys, cost analysis, values analysis, policy analysis, public speaking, writing, editing, computers, communications technology, and project management" (2002). King recommends four types of members to be part of this group including "evaluation

champions who will supply ongoing commitment, enthusiasm and technical know-how,” “highly respected staff who know the organization’s people and culture well, “competent people who can get things done,” and “At least one person with a sense of humor and the ability to keep the evaluation capacity-building process in historical context...” (2007, p. 50).

The integration of evaluation into the organization’s policies and procedures (Volkov & King, 2007) helps give evaluation a seat at the table and moves it toward being mainstreamed in the culture as a basic assumption. It also contributes to consistent and cohesive approaches to program evaluation (Milstein, Chapel, Wetterhall & Cotton, 2002). ECB involves the building of an infrastructure to support evaluation by facilitating learning and communication. Structures for learning may include training and coaching, access to evaluation learning materials, ongoing learning activities, and opportunities for reflective discussion about evaluations (Preskill & Russ-Eft, 2005; Preskill & Torres, 1999; Stufflebeam, 2002; Volkov & King, 2007). One example of an evaluation learning material is *Building Evaluation Capacity* by Preskill and Russ-Eft that contains 72 training activities for stakeholders to learn holistically about the design and implementation of evaluations by covering issues around politics, ethics and culture (2005). Communication systems should include feedback loops, a monitoring and tracking system, and distribution channels for disseminating information about evaluation processes and findings (Preskill & Russ-Eft, 2005; Stufflebeam, 2002; Volkov & King, 2007).

ECB is driven by context (Baizerman, Compton & Stockdill, 2002; Bamberger, Rugh & Mabry; Fetterman, 2005; Patton, 1997; Stufflebeam, 2002; Volkov & King,

2007), so organizations interested in building capacity for program evaluation should recognize their internal and external *organizational context*. There are many strategies for addressing an organization's internal context's for ECB. Volkov and King suggest an overarching, balanced approach of building an internal supportive culture for ECB while integrating the demands from external stakeholders (2007) such as funding organizations, government agencies and accreditation organizations. Indicators of a supportive culture for evaluation may include "stakeholders' buy-in, participation, and support from all levels" (Stufflebeam, 2002). This support would be necessary for mainstreaming evaluation in the organization to where it became viewed by stakeholders as part of what the organization does (Sanders, 2003), possibly to the level of other operational basic assumptions such as fundraising, financial management and human resource management.

Organizational culture is an important factor for the internal context for ECB and participation in program evaluation (Baizerman, Compton & Stockdill, 2002; Grudens-Schuck, 2003; Marais, 1998; Mesch & McClelland, 2006; Poole, Davis, Reisman & Nelson, 2001). An example of how culture can play into ECB is when an organization makes program evaluation a priority by conveying its importance to key stakeholders and by participating in the process at some level (United Way of America, 1996). Leadership can be the driver of this process internally by helping "...staff to not see this as an 'add on' in job tasks – but, instead, as an activity that is part of their essential job duties and part of the core responsibility of the organization" (Mesch & McClelland, 2006) (see *Leadership* section).

ECB practitioners desiring to help develop and nurture this type of organizational culture should have a pulse of the organization's readiness and willingness for change, social norms and customs, history of the organization and its traditions, work and management styles, relationships and power, and the overall stability of the organization. If ECB practitioners, such as evaluators who may facilitate this cultural change, do not recognize these factors, cultural change will be extremely difficult. Cultural change is also difficult when one person is trying to influence the values, beliefs, and norms of an entire group of people so the continuing theme of stakeholder involvement applies here as well. Allowing opportunities for participation and input for decision-making can be a powerful lever in the attempt to change the culture to one of interest and support for evaluation. However, if this effort compromises ethics, standards or guiding principles, the ECB practitioner may have no choice but to retreat from this effort.

NHSOs are open systems that rely on resources from the environment to operate and ultimately survive and this intensifies the power of resource dependency on the organization's decisions. In addition to the acquisition of resources, NHSOs' decisions are also influenced by the political environment particularly with regard to regulatory laws, government contracts and the changes in priorities and funding that go along with changes in administrations. In an even broader sense, NHSOs need to be cognizant of the current societal values and economic conditions, and how they affect their organization. If an organization allows these external environmental factors to dominate its decisions and actions, it runs the risk of drifting from their mission, losing their identity, losing their autonomy, and possibly losing the public trust.

These external environmental factors cannot be ignored, but at the same time researchers and scholars suggest that they be recognized, understood and integrated into the organization's participation in the evaluation and ECB processes. For example, they recommend that an organization become knowledgeable about its external environment, identify external mandates and utilize their influence to help build the culture for evaluation through innovation, accreditation, requirements for funding, innovative practices, other accountability demands and external support for evaluation (Stufflebeam, 2002; Volkov & King, 2007). Baizerman, Compton and Stockdill have observed that ECB is most effective when the demand and purpose for the evaluation are aligned (2002). This alignment requires a clear, transparent process for the stakeholders so they can determine whether the purpose of the evaluation is for program improvement and learning, which the ECB literature emphasizes, solely for accountability or a mix of both (Baizerman, Compton & Stockdill, 2002) which would represent an integrative, balanced approach that Volkov & King advocate.

Approaches

Baizerman, Stockdill and Compton tell us that ECB is emerging as a field of practice and that it is part art, craft and science (2002, p. 113). They indicate that reflective practice is at the core of ECB because it is necessary for practitioners to simultaneously be mindful of the big picture and the details. It is also necessary to build the knowledge base and peer learning community for ECB, similarly to the reflective practice of human service professionals (see page 99). This section will discuss various approaches found in the literature that contribute to ECB.

NHSO stakeholders and evaluators of their programs share the challenges of balancing the feasibility and the cost of evaluating programs with maximizing the provision of relevant, meaningful, accurate and useful information about those programs. *Evaluability assessments* were developed in the 1970's but they have recently been an emerging tool that nonprofit organizations can utilize to determine their capacity to implement a meaningful and useful evaluation of their program(s). An evaluability assessment is described as a process conducted before the program evaluation that ensures the program designs, the reality of the program's capabilities and other aspects of the program are ready for summative evaluation (Grinnell, Jr. & Unrau, 2005; Trevisan, 2007; Wholey, 1994). They help determine whether programs have the conditions necessary to be evaluated (Monette, Sullivan & DeJong, 2002; Patton, 1997; Trevisan, 2007; Wholey, 1994). Wholey, offers the following criteria for evaluability assessments:

“Program goals, objectives, important side effects, and priority information needs are well defined.
Program goals and objectives are plausible.
Relevant performance data can be attained.
The intended users of the evaluation results have agreed on how they will use the information.” (Wholey, 1994, p. 16).

Trevisan's study of published literature on evaluability assessments from 1986-2006 revealed that the three most common methods used were interviews, document reviews and analyses, and site visits (Trevisan, 2007, p. 295). It also revealed the top two reasons for conducting these assessments were to “assess program readiness for impact assessment” and to conduct “formative evaluation” (Trevisan, 2007, p. 296).

The benefits from conducting evaluability assessments include, but are not limited to, reducing costs associated with the evaluation; ensuring the evaluation produces results that are relevant, meaningful and useful; and saving stakeholder time spent on the

evaluation. NHSOs, who are interested in formally evaluating their programs for the first time or who are attempting to satisfy new demands from external stakeholders, run several risks of moving forward with evaluating their programs without first conducting an evaluability assessment. They may include realizing after the fact that their staff lacks the appropriate knowledge to properly work with the evaluator, the program objectives were ill-defined, the organization does not have the data collection infrastructure and processes in place necessary to collect the relevant data, and other related issues. Ultimately, evaluability assessments help the stakeholders involved in program evaluation better understand the expectations of those holding influence over the program and their differences; explore costs, feasibility and utility of the proposed evaluation and promote the use of evaluation results (Wholey, 1994).

Patton suggests that when evaluators conduct evaluability assessments they are really engaging in program and organizational development (1997, p. 104). This is an important point because just as evaluability assessments help stakeholders determine a program's readiness for evaluation, they also contribute to determining the organization's readiness to engage in program evaluation. There may be situation where an organization appears ready for program evaluation based on the evaluator's assessment of their overall evaluation capacity, but the program in question is not ready due to unspecified goals, measurements or other aspects. Conversely, the program may be sound and ready to be evaluated, but the evaluator determines that the organization is not ready to move forward based on a lack of resources, unsupportive culture, or other factors.

Preskill and Torres developed a tool that can be used to measure evaluation readiness on an organizational level called "The Readiness for Organizational Learning

and Evaluation (ROLE) instrument”⁶ (Russ-Eft & Preskill, 2001). The ROLE instrument is a survey containing 78 items, mostly Likert scale, grouped in six areas of scope - Culture, Leadership, Systems and Structures, Communication, Teams and Evaluation. It can be administered to various stakeholders, typically the organization’s staff, to get either a departmental perspective or one for the entire organization. The aggregate results of the surveys can help an organization determine the areas where it has a high level of readiness for organizational learning and evaluation and the areas that need to be shored up before the organization proceeds with evaluating its programs.

Combining evaluability assessments on program and organizational levels can help determine the overall readiness for a NHSO’s program to be evaluated by analyzing the conditions for evaluation. Patton offers his wish list for ideal evaluation conditions that includes items like clear, specific, and measurable program goals, formative and summative evaluation, dedicated staff, stakeholder enthusiasm and cooperation, enough resources and time, and the fantasy of having no surprises throughout the process among others (1997, p. 118).

Stakeholder involvement is a dominant theme throughout ECB. One example of an approach that impacts ECB driven by such involvement is *participatory evaluation* which is defined as, “evaluation intended not only to improve program understanding but also to transform program-related working relationships through participation in evaluation” (Greene, 1997 in Bamberger, Rugh & Mabry, 2006). Cousins and Whitmore’s describe it as “a practical approach to broadening decision making and problem solving through systematic inquiry” (1998, p. 87). King argues that all

evaluation is participatory because by default it involves some level of interaction between the evaluator and other stakeholders (2004, p. 337).

Participatory evaluation lends itself well to ECB because it involves experiential learning. It embodies the constructivist's perspective that people bring perceptions to issues based on their past experiences and gain knowledge and new understandings for these issues through participating in an active learning process. Somers indicates that this main feature makes participatory evaluation is a good choice for informal educational settings (2005). They would comprise non-school learning opportunities such as after-school programs, adult education programs, youth educational programs such as those delivered by the scouting organizations and Junior Achievement, and environmental education programs, all of which have a prominent place in human services. Somers also indicates that participatory evaluation is advantageous for ECB in these settings because they are in great need for evaluation capacity at the local level, and stakeholders usually lack evaluative skills and knowledge (2005). Other advantages this approach has for ECB are that it is cost-effective in lieu of a more expensive formal training program, it is focused on each individual's needs and how program evaluation relates to their specific responsibilities, and it is specific to the situation, organization and time for the evaluation process (Somers, 2005).

Participatory evaluation is not without challenges that are to be expected for such a democratic group process. The stakeholders must address issues of power, ethics, selection, quality, culture, training and others (Cousins & Whitmore, 1998, p. 101). Evaluators as facilitators, conveners, coordinators and managers have the challenge of using their discretion when addressing these issues regarding their level involvement in

the evaluation process, individual and organizational boundaries and what might constitute an infringement on those boundaries. King so reminds us of what I call the evaluator's serenity prayer by stating that even when evaluators take responsibility for evaluation use, they ultimately can't force anyone to use the results. A helpful, if not therapeutic, reflective exercise for evaluators is to realize what they can and cannot control, and have the wisdom to know the difference.

A more intense version of participatory evaluation is *empowerment evaluation* where stakeholders are more than just involved in the evaluation process; they take control of the process. Other approaches such as participatory evaluation or collaborative evaluation might result in stakeholder empowerment, but in empowerment evaluation, it is a primary objective. Empowerment evaluation is defined as:

“An evaluation approach that aims to increase the probability of achieving program success by (1) providing stakeholders with tools for assessing the planning, implementation, and self-evaluation of their program, and (2) mainstreaming evaluation as part of planning and management of the program/organization” (Wandersman, Snell-Johns, Lentz, Fetterman, Keener, Livet, et al., 2005, p. 30).

While other approaches focus more on utilization, empowerment evaluation focuses more on the theme of social justice. In spite of this main focus, we can infer that empowerment evaluation incorporates utilization from the reference to “mainstreaming” and the principles it is based on. The principles for empowerment evaluation include *improvement, community ownership, inclusion, democratic participation, social justice, community knowledge, evidence-based strategies, capacity building, organizational learning, and accountability* (Wandersman, Snell-Johns, Lentz, Fetterman, Keener, Livet, et al., 2005, p. 30). These principles are compatible with evidence-based practice and the standards for the human service professions of social work, workforce development,

psychology, counseling and others. This compatibility makes empowerment evaluation a palatable choice for ECB in NHSOs.

The California Department of Health Services Tobacco Control Program (TCP), while not an example of a NHSO, serves as a good example of empowerment evaluation in action and how it can benefit an organization and its program. The benefits TCP realized, some of which are referenced here, came from lessons they learned throughout the evaluation process. They decentralized their program evaluation resulting in a balance between using uniform measures for consistency while also customizing evaluation for local program issues. In similar fashion, they revised their training of their funded agencies to include evaluation designs in more detail that match more closely the local programs' objectives. Lastly, TCP worked more closely with their funded agencies by improving communication, providing information on how to select an evaluator, providing training for program staff and local evaluators, and convening a work group comprised of a mix of program staff and evaluators (Tang, Cowling, Koumjian, Roesler, Lloyd & Rogers, 2002). Overall, TCP's experience with empowerment evaluation indicated that program evaluation helped create knowledge for stakeholders and build evaluation capacity for their organization and their funded agencies.

Bamberger, Rugh and Mabry's *RealWorld Evaluation Capacity Building* (RWE) emphasizes some common characteristics found in other approaches for ECB such as stakeholders gaining knowledge, acquiring skills, and developing an understanding for and willingness to use evaluation (2006). RWE capacity building differentiates from these approaches by adding the practical countenance of demonstrating that evaluations can maintain quality and rigor while dealing with "real-world constraints" such as

resources, time, etc. (Bamberger, Rugh & Mabry, 2006, p. 357). RWE capacity building also extends its skill-building objectives beyond the program staff or other stakeholders in the organization running the program. It suggests that capacity building also strengthen evaluation skills of other groups who support the evaluation by tailoring the approach based on their specific learning needs. They include organizations that fund or commission evaluations; evaluation practitioners; evaluation users; stakeholder groups impacted by the evaluation such as community groups or trade associations; and the general public (Bamberger, Rugh & Mabry, 2006, pp. 358-359).

The categories of knowledge and skills RWE capacity building addresses include “Defining evaluation needs and commissioning evaluations, designing evaluations, implementing the evaluation, data analysis, disseminating and using evaluations and conducting evaluations under real-world constraints” (Bamberger, Rugh & Mabry, pp. 359-360). Some of the specific items within these categories include determining the resources and technical requirements for the evaluation, conducting an evaluability assessment, establishing program theory, developing evaluation’s language (terms of reference), supporting utilization, and incorporating methodology that accommodates real-world constraints while at the same time maintains validity and overall quality (Bamberger, Rugh & Mabry, 2006). RWE capacity building incorporates a broad, holistic approach to ECB by going beyond involving stakeholders to ensuring they have the appropriate knowledge and skills necessary to fulfill their role in the evaluation process.

We might hypothesize that many of the documented challenges and failures of program evaluation have resulted from secondary or tertiary stakeholder groups not

having the proper understanding of evaluation and not being equipped to fulfill their role. For example, it is interesting to note that the RWE approach includes funders implying that their role goes beyond just requiring evaluation and funding it to understanding the evaluation's needs and what the evaluation can and cannot do. Involving the funders in this ECB approach is intended to equip them with what they need to effectively engage in the evaluation process, but it might have additional effects such as helping balance the imbalanced power relationship between them and NHSOs through dialogue, responsibility and a sense of ownership.

Appreciative inquiry (AI) is defined as “a group process that inquires into, identifies, and further develops the best of ‘what is’ in organizations to create a better future” (Preskill & Catsambas, 2006) and has been described as “a collaborative and highly participative, system wide approach to seeking, identifying, and enhancing the ‘life giving forces’ that are present when a system is performing optimally in human, economic, and organizational terms” (Watkins & Mohr in Elleven, 2007, p. 451). AI is based on an examination of what is currently working in an organization, at times what is working well, and using that information as the foundation from which to move forward and get to the desired state or condition. It is driven by some common themes in other ECB approaches such as stakeholder involvement, reflection, learning, action and change. Some suggest that the key to transformative change is AI's provision of opportunities for active reflection (Donovan, Meyer & Fitzgerald, 2007).

AI has been used for a myriad of purposes usually within the frame of organizational change such as for developing leaders, redesigning organizations, planning the future of an organization in settings such as workshops, seminars, and consulting

projects. AI is similar to the asset-based approach the Search Institute uses for their 40 Development Assets for Youth. Instead of taking the classic social science approach of identifying problems youth encounter and determine how to treat them, they examine the conditions, behaviors and actions that contribute to youth growing up healthy, happy and confident (Search Institute, 2007). Preskill and Catsambas state that “Appreciative Inquiry is about recognizing the best in people; acknowledging those things that give life; affirming past and present strengths, successes, assets, and potentials; and asking questions, studying, and searching, exploring, and investigating” (2006, p. 3). They apply AI to ECB by confronting what might be limited, negative, mythical or stereotyped views of evaluation that stakeholders may have. These stakeholders also may view their involvement in the evaluation process as an extra strain on their already stretched time, so AI can help them see the benefits are worth their investment in time. AI can help transform these views by having them see evaluation as relevant to them and their organization, an important learning process, and the basis for integrating their past experiences into the collective body of knowledge about their programs.

Preskill and Catsambas suggest AI’s role in ECB is to enhance the facilitation, communication, understanding and support for evaluation (2006). They offer seven specific ways how AI accomplishes this:

1. Reframes the study of problems to the study of successes
2. Emphasizes how evaluation can be a learning process rather than a punitive process
3. Provides an option for more cost-effective evaluation
4. Contributes to culturally responsive evaluation by embracing diversity
5. Offers new language that allows greater honesty about difficult topics
6. Unleashes creativity through affirming, participatory, and energizing processes
7. Increases understanding of evaluation processes and findings, thus leading to greater use and influence

8. Complements systems thinking and complexity theory approaches (Preskill & Catsambas, 2006, pp. 132-137).

Not all researchers and scholars agree that program staff and administrators of NHSOs should have to gain evaluation skills for the ECB process. Miller, Kobayashi and Noble offer *insourcing* as an alternative to more commonly known ECB approaches which typically either target internal stakeholders building skills and capacity to conduct their own evaluations or outsourcing which involves contracting with an external evaluator (2006). They describe insourcing as a hybrid of these two approaches that has the “intention to optimize the partnership between evaluators and program staff” (Miller, Kobayashi & Noble, 2006, p. 86). It is a process where these two stakeholders meet in the middle, with the program staff committing to acquire a working understanding of evaluation and be motivated to help in the process, and the evaluator committing to developing a structure and process that minimizes the burden for staff. Each has specific roles and responsibilities in the evaluation process with little overlap between them.

Insourcing differs from other ECB approaches in that 1) it does not have the goal of building evaluation skills for staff and 2) it is a minimalist approach as opposed to more comprehensive or holistic approaches. The overall theme of their approach is couched as a compromise for enabling evaluation to be less expensive and sustained in the organization, or as they describe it – “evaluation with humility” (Miller, Kobayashi & Noble, 2006, p. 93). The message conveyed is that such compromising practical evaluation is better than a comprehensive ECB effort that is not realistically sustainable. As a result, insourcing is geared to smaller organizations, which comprise the majority of human services organizations. The authors caution funders to not use insourcing as an excuse to reduce financial support for program evaluation. To the contrary, if evaluation

is to sustainable it requires continual support, as in the case of the programs being evaluated.

Patton's definition of *process use* provides yet another approach that can contribute to ECB:

“Process use refers to and is indicated by individual changes in thinking and behavior, and program or organizational changes in procedures and culture, that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process” (1997, p. 90).

Process use can contribute to ECB in a more action-oriented manner by expanding an organization's readiness to conduct and use evaluation through experiential learning (Amo & Cousins, 2007). The changes in attitudes, behavior, and the gaining of knowledge and skills through process use can also change the culture for evaluation when stakeholders internalize the logic for evaluation, and work towards institutionalizing it by incorporating it as a component of running the program (Fetterman, 2003).

King points out that for process use to have these effects it should not be left up to the possibility that it will likely influence stakeholders in this way. For it to have a chance of reaping these results, it must be intentional and have someone driving process use towards ECB (2007). This intention can take the form of an evaluator as an educator, as in developmental evaluation, who ensures process results in stakeholder learning. Learning through process use comes from testing, experimenting and realizing the challenges for program evaluation based on a given organizational context and developing strategies for overcoming them.

King suggests some strategies for evaluators desiring to use process use for ECB. One is to assess the contextual and cultural viability for ECB to determine whether intentions to apply process use in the organization have a chance for success. Identifying

stakeholders who sincerely care about evaluation and its findings (champions), and supporting them, will be necessary for these champions to nurture the same feelings in others and eliminate the chance the evaluator is alone in the effort for cultural and organizational change. The evaluator in order to be intentional should be resolute, upbeat, and inclusive when negative perceptions or feelings are expressed that may hinder the process. Lastly, the evaluator needs to “work with evaluation champions to construct an evaluation process and structures to support continuing evaluative thinking in the organization” (King, 2007, p. 52).

This section has attempted to provide a brief and broad overview of the constructs and approaches for ECB to establish the concept for which this research is analyzing⁷. Some common themes throughout this discussion include stakeholder involvement, reflection, learning, contributions to use, change, and improvement. This section also reinforces that ECB is reliant on context, resources, structures, and culture. These actions and concepts parallel those in the human service professions.

The following sections intend to broaden our scope for ECB and its relationship to human services by providing background for issues specific to the world of NHSOs that have a role in ECB. They include nonprofit management; the role of a NHSO’s mission; NHSOs activity in program evaluation; a more detailed account of organizational culture’s role; the impact that leadership can have on organizational culture; the importance of managing the socio-political environment for the ECB process; choices NHSOs and their leaders can make and actions they can take in responding to the challenges for program evaluation; the relationship of ECB and organizational learning;

and the roles of the five stakeholder groups included in this research – EDs, board chairs, program staff, funders and evaluators.

Nonprofit Management

Program evaluation has been increasingly found in the nonprofit management literature over the past decade (Campbell, 2002; Fine, Thayer & Coghlan, 2000; Hoefer, 2000; Oster, 1995; Paddock, 2001; Poole, Davis, Reisman & Nelson, 2001; Smith, Bucklin & Associates, 2000; Thomas, 2005; United Way of America, 1996). Various reasons are cited for why it is an important component of nonprofit management for nonprofit organizations and their stakeholders. Program evaluation helps organizations determine whether their program has met its objectives and achieved desired results (Paddock, 2001; Smith, Bucklin & Associates, 2000; Thomas, 2005); provides a basis for comparison against similar modes of service delivery or industry “best practices” (Paddock, 2001); serves as a tool for making decisions about programs (Paddock, 2001; Smith, Bucklin & Associates, 2000; United Way of America, 1996); and enables an organization to meet the demands of its funders and other external stakeholders (Hoefer, 2000; Thomas, 2005). The ultimate responsibility for program evaluation in a nonprofit organization resides with the ED (Thomas, 2005), as he or she is in the position to ensure that it is planned and executed with the proper levels of participation from key stakeholders, especially staff. Stakeholder involvement is important as it “...increases the likelihood that evaluation results will be used and that evaluation processes will continue” (Fine, Thayer & Coghlan, 2000; p. 334). The responsibility also resides with the board of directors who are responsible for ensuring that programs are meeting the

needs of the constituency being served and that they are contributing to work towards the organization's mission (Burgess, 1993). The roles of the ED and the Board of Directors will be discussed more specifically in their stakeholder roles sections.

One study revealed that resources such as funds, staff time and expertise were the top three reasons why programs were not evaluated (Hoefer, 2000). This suggests that program evaluation ultimately relies on the ED and the board who play primary roles in the acquisition, allocation and expenditure of resources through planning, fundraising, and budgeting. Some suggest that the ED's role is critical to the success of an organization's program evaluation process through ensuring it's a priority within the organization, modeling behavior by participating in the process and communicating evaluation information to the board (Thomas, 2005; United Way of America, 1996). We can logically conclude that the program evaluation process represents a management tool that can be initiated, fostered and implemented by an organization's leadership, as it requires the management functions of planning, budgeting, organizing, staffing, controlling and problem solving.

The Role of Mission

The mission of a 501 (C) (3) nonprofit organization embodies its primary (charitable) purpose or reason for being (Bryson, 2005; Dym & Hutson, 2005; Gibelman, 2003; Nanus & Dobbs, 1993; Poister, 2003; Stern, 1999). This primary, charitable purpose allows it to attain tax-exempt status by the Internal Revenue Service (IRS) and places it in the public's trust to carry out that purpose. The mission represents an entrepreneurial idea (Oster, 1995), and in NHSOs is typically developed for the purpose of addressing the unmet or insufficiently met human needs of individuals, families and

groups of people in their community and achieving specific ends (Gibelman, 2003). It serves as the foundation for stakeholders to rally around as well as a guideline for how the organization will serve the public good (Minkoff & Powell, 2006). The mission is expected to represent the organization's values, philosophy and ethical standards (Bryson, 2005; Gardner, 2006).

The intended results of an organization's program(s) evolve out of the mission's development (Bryson, 2005; Stern, 1999), as the mission serves as the basis for developing strategies, meeting objectives and measuring performance (Dym & Hutson, 2005; Hudson, 2005; Poister, 2003; Werther & Berman, 2001). Nanus and Dobbs also characterize the mission as "the maximization of the social goods they produce for both society and the people who participate in them" and "the single most important measure of success of nonprofit organizations" (1993, p. 39). In NHSOs, programs serve as vehicles through which the mission is carried out (Gibelman, 2003). A properly developed mission sets up an evaluative framework connected by the program's outcomes by determining the aspects of the program(s) that are to be assessed (Monette, Sullivan & DeJong, 2002). These aspects and their corresponding outcomes are driven by the goals of the program which are developed from the mission. The developmental relationship between the mission of a NHSO and program evaluation are depicted below in Figure 4.

Figure 4 – Relationship Between Mission and Program Evaluation

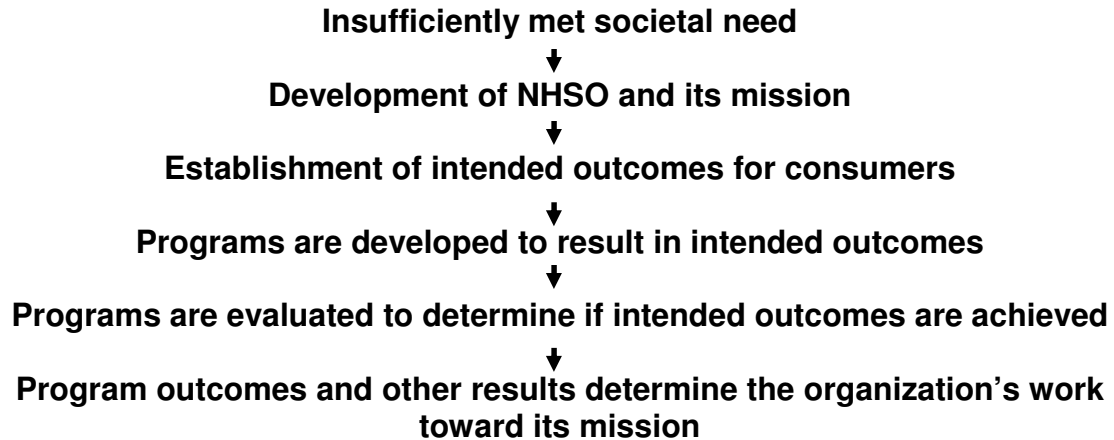


Figure 4 displays a direct link between an NHSO's mission and program evaluation through the effort to find out how well the organization is serving its consumers and contributing to its mission. Therefore, we can conclude that program evaluation in NHSOs serves as the means for assessing the organization's work towards their mission.

The mission also serves as an expression of the organization's culture (Bjerke, 1999). As a result, it can effectively serve as a "litmus test" for an organization's activities and a link between culture and strategy. Leaders can use the mission as the basis for decision making and justifying organizational change. If leaders comprehend and value the proposed relationship shown in Figure 4, they can use the mission as the basis for embedding program evaluation in the organization's culture as a basic assumption. Organizational change resulting from program evaluation can be less resisted by stakeholders if it is culturally viewed as change to improve work towards the mission. This mindset can move stakeholders towards understanding and valuing program evaluation as an important tool for assessing the effectiveness of how the organization serves its consumers.

NHSOs and Program Evaluation

Human service programs are often designed based on assumptions for the causes of social problems and how to address them. The assumptions are not always explicitly stated by NHSOs and the degree to which these assumptions are based on prior information varies among organizations. They can be based on a single or combination of sources including the current human service literature on the particular social problem being addressed, the latest research on the social problem such as social indicators, recommended practices from the particular profession dealing with the social problem, and/or prior evaluation studies conducted for the program or similar programs in other organizations (Martin & Kettner, 1996). Some suggest it is natural to couple human services with program evaluation because of the field's continual quest to answer questions about human behavior, social problems, and service delivery (Monette, Sullivan & DeJong, 2002). Human service activities also have a purpose specific to the social problem or problems being addressed. Program evaluation enables NHSOs to check against and update programs' assumptions (Underwood & Lee, 2004) as well as determine if the purpose of their activities yield intended results (Thompson, 2000) by demonstrating the program's effectiveness.

Program evaluation can be a valuable tool for NHSOs to focus on the needs of their consumers, respond to external stakeholder demands, adapt to changes in public policy, tell their story in a more meaningful way, ensure quality control, remain competitive, and support their advocacy efforts. It can also enable a NHSO to deliver their programs within an environment and culture for continuous improvement driven by their mission as depicted in Figure 5 below.

Figure 5 – Program Improvement Continuum



The effectiveness perspective moves the criteria for program success beyond efficiency which tends to measure success by how resources have been used and how many people have been served. This perspective with regard to accountability is concerned with maximizing outcomes with inputs, in other words successfully impacting the lives of consumers with the currently available resources (Martin & Kettner, 1996). Program evaluation in the context of a NHSO serves to demonstrate how effective the organization's programs are at intervening in the lives of its human consumers to address their social needs. The essence of accountability in the rationale for human service practice "clarifies the basis of the intervention and the objectives set; explains the actions taken to meet those objectives and the reasons for doing so; evaluates the intervention" (Thompson, 2000, p. 35).

Royse reminds us that “Ultimately, program evaluation benefits the clients. It tells us whether clients are being helped or not, and it can indicate how we can better assist clients” (1991, p. 194). This is congruent with the primary objective of, and often the motivation for, the work of program staff. However some evaluation activities, such as reporting and other administrative functions, may be viewed as additional responsibilities in a profession known for stretching people to their limits and one that created the phrase “burn out.” These demands, typically from external stakeholders, can be counterproductive to an organization’s mission and the social worker’s desire to serve their consumers (Christensen & Ebrahim, 2006). How program evaluation tasks are integrated in program staff’s responsibilities can impact their perspective and support for it. In a study of NHSOs serving immigrants and refugees by Christensen and Ebrahim, program staff saw needs assessments and evaluations, “...as fundamental to serving clients. Rather than thinking of their activities in terms of accountability, staff members connect these actions to doing their jobs well...and view them as...necessary to know that they are meeting client needs...” (2006, p. 205).

Studies that examine the motivations behind NHSOs engaging in program evaluation are not common; however, we can see the mix of environmental and internal forces in their results. For example, in Fine, Thayer, and Coghlan’s study of 140 nonprofit service delivery organizations the top two most frequently cited reasons (with a respondent’s option to select more than one) for evaluating their programs were to measure impact or outcomes of the program at 56%, and satisfy a funding requirement at 43% (2000, p. 333). Hoefer’s study of 91 NHSOs in the Dallas, Texas area revealed that the top three most frequently cited reasons (with a respondent’s option to select more

than one) were compliance with procedures at 57%; curious to see how their program was doing at 51%; and demonstrating the value of their program to funders at 42% (2000, p. 171). He also examined their reasons for not evaluating their programs. Not having enough funds and not having staff available for the evaluation were the two top reasons tied with 48%. The second most frequent reason was because the funder did not require it with 43%. Carman's study of approximately 100 NHSOs in New York State showed that 59% indicated the most descriptive statement of why their organization evaluates its program(s) was "It helps us to gather information about our programs, make decisions and improve our programs" with the second most descriptive statement, "We are required to do program evaluation (by funders, the board or management)" at 16% of respondents (2005, p. 111).

Carman also conducted a study to find out how community-based organizations (CBOs) are responding to the demand for program evaluation, specifically inquiring about the activities they engage in for evaluation, the types data collected, their method for collecting that data, who is primarily responsible for the evaluation activities and the sources of funding that pay for the evaluation (2007, p. 62). Approximately 80% of the 178 organizations of varying sizes reported that they engaged in program related activities such as reviewing information, monitoring its implementation and determining whether it has met its goals and objectives. This portrays their activities to be more related to program information compliance rather than what the literature defines as program evaluation.

About two thirds of the organizations gathered some kind of data on their program's results or outcomes. There was variance among the types of CBOs for what

they considered and how they characterized outcomes representing an overall challenge for outcome measurement acknowledged by most of their executive directors (Carman, 2007). Other challenges for measuring program outcomes were more specific to the service delivery type, as in prevention programs where outcomes comprise what doesn't happen.

The most common types of methods for collecting program evaluation data were written tools such as surveys and intake forms but they were used primarily for gathering service delivery information as opposed to gathering evaluation data. This most likely was reflected by the fact that many characterized program evaluation as compliance or quality assurance. The overwhelming majority of the organizations reported that "executive or management staff" was responsible for the evaluation and not surprisingly, internal evaluators on staff were rare (Carman, 2007). Internal operating funds were the most frequently referenced source of funds for the evaluation process possibly indicating a lack of financial support for evaluation from funders, some of whom are asking them for evaluation information.

Carman's conclusions present implications for ECB in NHSOs. First, "program evaluation" and "outcomes" were characterized differently among organizations and often related activities such as monitoring or reporting were used to what they considered evaluation. A common theme through the ECB literature is the importance of stakeholders having a clear and consistent understanding of program evaluation so the evaluation process can run efficiently and effectively, and program evaluation can be meaningful to them. This is usually one of the first steps in the ECB process. If

Carman's study were to be representative of all NHSOs throughout the U.S., there is plenty of technical assistance and basic ECB work to be done.

Another implication is an overall lack of organizational investment for evaluation with regard to the data they're collecting and the infrequent use of available evaluation tools. For example, logic models were rarely used possibly indicating that the program's theory has not yet been formally established. Funding for evaluation to either hire an internal evaluator or contract with an external one is scarce. This combination of lacking an understanding for evaluation is, overall expertise and financial support is likely driving this scant level of investment and evaluation capacity.

Carman recommends that funders can help break these mindsets of their grantees by beginning to ask for data that demonstrates how effective their programs are and what efforts have been made to improve them as opposed to perpetuating the monitoring or quality assurance mentality (2007). She also suggests that evaluators can step up to the plate and help fill the knowledge and skills gaps in these CBOs. They could also seek strategies for reducing costs and show them how to conduct low-cost evaluations. Lastly, Carman calls on CBOs to invest in evaluation the same way they would for other management functions and also build capacity to take advantage of time and cost-saving technological tools and methods, such as web-based surveys for example.

The infusion, and increasing use, of outcome measurement has been a driving force behind the evaluation of nonprofit human service programs over the past few decades. Plantz, Greenway and Hendricks emphasize that "the most important reason for implementing outcome measurement is that it helps programs improve services..." and that "...its value in enhancing service effectiveness should be seen as primary" (1997,

p. 23). The United Way of America predicted in 1996 that “Outcome measurement is not a passing fad. In the years ahead, collecting data on benefits for program participants will be as common as collecting data on the number of program participants is today” (1996, p. 8). There is no solid evidence twelve years later to confirm their prediction; however United Way’s role in the outcome measurement movement that began in 1996 appeared to have effects on NHSO activity. Plantz, Greenway and Hendricks reported several national outcome studies conducted by Big Brothers Big Sisters of America, The Child Welfare League of America, Girl Scouts of the U.S.A., and Girls, Incorporated, as well as several examples of national offices of NHSOs attempting to build evaluation capacity on a local level by providing tools or resources to their affiliates (1997, pp. 20-21).

Outcomes for human service programs demonstrate the quality of life changes in consumers as results, impacts or accomplishments (Martin & Kettner, 1996). These changes may include gaining knowledge, and/or changes in condition, status, or behavior resulting from experiencing a program’s intervention. An example of gaining knowledge would be parents learning new techniques for disciplining their children from parenting workshops. A change in condition might be moving from homelessness to transitional housing. The status of a consumer can be changed if he or she acquires employment. A client can change their behavior through rehabilitation from alcoholism.

Outcomes are specific to the social issue being addressed and the organization’s lens for framing that issue, the program’s service delivery framework, and the characteristics of the consumers being served. These factors are not static, as the environment in which NHSOs operate is constantly changing, and therefore programs

must be flexible to address these changes and the needs of their consumers.

Polkingborne suggests that “...it holds that better outcomes are produced by adjusting practitioner actions to the characteristics of specific situations and to the changes taking place in the individuals being served” (2004, p. 4). Within this context, outcomes are logically connected to the program as the results from services being delivered to the consumer (Schalock, 1995). For example, we would expect a service learning program for youth to include intended outcomes linked to the learning objectives of the service in which they engage. Similarly, a smoking cessation program would intend to produce the outcome of participants stopping smoking. One of the pitfalls for selecting outcome measures is not linking them to the organization’s mission by not remaining focused on the specific social problem being addressed (Martin & Kettner, 1996).

The assessment of a program’s intervention remains a primary reason for the evaluation of social work (Reamer, 1998). Another important reason for program evaluation is its impact on the practice of human service professionals. First, it helps them determine the effectiveness of their practice (Eliason, 2007; Gardner, 2006; Reamer, 1998; Schuerman, 1983). Scholars that make this claim do not portray program evaluation as an “administrative task” or “add-on” process. For example, Schuerman suggests it gets at the essence of practice by asking what activities are effective under what circumstances (1983, p. 6). Human service professionals are charged with identifying the indicators of this effectiveness through gathering, analyzing and interpreting data (Reamer, 1998). Second, these professions, such as social work, have a foundation of knowledge that evolves and informs its practice. Human service practitioners are expected to continue their professional development, learn from their

experience and be knowledgeable on the current research and evidence-based practice (Eliason, 2007; Gardner, 2006; Thompson, 2000). Program evaluation enables practitioners to contribute their feedback to the profession's practice-based research (Eliason, 2007) and make knowledgeable choices for interventions to be used with their consumers.

The National Association of Social Workers' (NASW) code of ethics states for the ethical responsibility of *Competence* that:

“Social workers should strive to become and remain proficient in professional practice and the performance of professional functions. Social workers should critically examine and keep current with emerging knowledge relevant to social work. Social workers should routinely review the professional literature and participate in continuing education relevant to social work practice and social work ethics. Social workers should base practice on recognized knowledge, including empirically based knowledge, relevant to social work and social work ethics.” (NASW, 1999).

NASW's ethical responsibility for *Education and Training* reads,

“Social workers who function as educators, field instructors for students, or trainers should provide instruction only within their areas of knowledge and competence and should provide instruction based on the most current information and knowledge available in the profession.” (NASW, 1999).

The ethical responsibility for *Integrity of the Profession* reads,

“Social workers should contribute to the knowledge base of social work and share with colleagues their knowledge related to practice, research, and ethics. Social workers should seek to contribute to the profession's literature and to share their knowledge at professional meetings and conferences” (NASW, 1999).

These standards help portray the process that leads to *research-based* or *research-minded practice* which combines intellectual inquiry with practical application and has been described by Everitt et al. as being “...concerned with the analytical assessment of social need and resources, and the development, implementation and evaluation of

strategies to meet that need” (Thompson, 2000, p. 61). Human service professionals develop this type of practice by integrating research and evaluation activities as part of their practice. Some of the goals of research-minded practice include balancing rigor with creativity (Thompson, 2000).

An example of the call for research-minded practice is found in the Code of Professional Ethics and Practices of the National Association of Workforce Development Professionals (NAWDP) for *Responsibilities to the Profession* which states that “Members contribute to the knowledge base of the workforce development profession by participating in and supporting research and other activities that identify successful strategies and programs” (NAWDP, 2002). Another example is found in the American Psychological Association’s (APA) ethical principle for *Use of Assessments* which reads, “Psychologists administer, adapt, score, interpret, or use assessment techniques, interviews, tests, or instruments in a manner and for purposes that are appropriate in light of the research on or evidence of the usefulness and proper application of the techniques” (APA, 2002).

Another term that involves human service practice based on knowledge and learning is *informed practice* which recognizes the important of incorporating theory into practice for an approach that is based on ‘lifelong learning’ driven by continuously engaging human service workers in professional development (Thompson, 2000). Thompson provides some common activities between theorists and practitioners to illustrate how the process of linking theory with practice can be accomplished:

- making sense of experience;
- making predictions/anticipating;
- relating events to a pre-existing body of knowledge;
- forming hypotheses and testing them out;

- using general principles to develop a framework of understanding;
 - addressing conflicts between one's own views and those of others.
- (2000, p. 136)

Reflective practice also joins theory with practice through reflective learning which involves human service professionals reflecting on their experiences, processing what they have learned and pondering their significance. This process is one the four stages in Kolb's four-stage cycle for experiential learning which includes moving from *concrete experience* to *observations and reflections* to *formation of concepts and generalizations* to *testing implications of concepts in new situations* (Kolb in Chickering, 1977). Reflective practice is important for human service professionals because they deal with uncertainty, their role makes them a part of the experience, and it helps their practice go beyond technical rationality by incorporating their perspectives and values (Gardner, 2006; Thompson, 2000). It can also be a morale boost for overworked professionals who can subconsciously adopt a day-to-day, 'put-out-fires' mentality towards their work and not take time to reflect critically on their work and the importance of their role in the impacting the lives of their consumers.

Evidence-based practice is rooted in the medical profession, but in the context of human services is defined as the professional's practice of integrating current best research evidence with their expertise to make decisions about how to most effectively address a human consumer's unique social problem(s). NASW offers the following definition,

"EBP is a process involving creating an answerable question based on a client or organizational need, locating the best available evidence to answer the question, evaluating the quality of the evidence as well as its applicability, applying the evidence, and evaluating the effectiveness and efficiency of the solution. EBP is a process in which the practitioner combines well-researched interventions with clinical experience, ethics,

client preferences, and culture to guide and inform the delivery of treatments and services” (NASW, 2007).

At first glance this term invokes an emphasis on the scientific evidence however; these definitions help remind us that this type of practice cannot be successful without trained and highly skilled human service professionals (Eliason, 2007). An example of the call for this balance comes from the American Psychological Association in their ethical principle for *Bases for Scientific and Professional Judgments* which asks that “Psychologists’ work is based upon established scientific and professional knowledge of the discipline” (APA, 2002). While there are a variety of descriptions for evidence-based practice, it is important to note that there is not a single superior approach to practice. (Eliason, 2007). Beyond the commonalities of incorporating the most current literature, research and evaluation into practice, different approaches will always be necessary to serve different social problems in different environment and contexts, and account for cultural competency within the various human service professions (Eliason, 2007).

Program evaluation seemingly has great potential to serve as one of the processes that contribute to research-minded, informed or evidence-based practice for program staff based on some common features and its ability to provide important information about how a program is delivered and the effectiveness of its intervention. In summary, program evaluation’s ability to provide information on consumer outcomes and the program’s process enable it to contribute to program theory and have an important role in the evolving research-practice collaborative effort in human services (Eliason, 2007).

Evaluation and human services delivery are both social interventions. In spite of that common characteristic, in the past evaluation was viewed by human service professionals as a process performed by people other than themselves, such as academic

researchers, but that perspective has evolved over time to becoming part of their practice and their toolkit of knowledge and skills (Reamer, 1998). Practitioners realize that issues of assessment and measurement are not just functions of researchers because, “Problems of validity, reliability and error can result in ineffective and possibly harmful practice intervention” (Monette, Sullivan & DeJong, 2002, p. 124). Before this evolution took hold in the human service professions, some reflection on our part would reveal that they have been incorporating evaluative skills all along including conducting needs assessments of their consumers, establishing and assessing goals, assessing the progress of their consumers, solving problems, building a knowledge base for their profession, and others. Reamer indicates that evaluation is the third goal that guides such work after conducting exploratory research to determine what has been done to address the particular social problem and acquire consumer feedback, and describing the social phenomenon being studied and the change in consumers over time (1998, pp. 18-21).

Evaluation is fundamental to the human service professions as it is the method for determining whether program staff’s interventions into consumers’ lives has resulted in the desired change. Beyond this evaluation of direct service to the consumer, the evaluation of human service programs can also contribute to assessing how they are addressing the particular social problem at the organizational, community and policy levels. Evaluation has now taken an official position in the principles and ethical standards for some human service professions and is becoming an integral part of all human service program staff work. This will be discussed along with the role of program staff in the program evaluation and evaluation capacity building processes in more detail in the stakeholder roles section for Program Staff.

Organizational Culture

The culture of an institution and its environment are important factors that drive individual and organizational behavior. Culture is described as “including ideas and beliefs, the affective/expressive dimension, and an evaluative element consisting of value-orientations” (Parsons in DiMaggio & Powell, 1991). Organizational culture, often referred to as corporate culture especially when referencing businesses, has been characterized as comprising shared values, ideas, beliefs, assumptions, norms, artifacts and/or patterns of behavior (Bjerke, 1999; Ivancevich, Szilagyi, Jr. & Wallace, Jr., 1977; Ott, 1989; Schein, 1992). Organizational culture has been more specifically defined as “...the importance for people of symbolism – rituals, myths, stories, and legends – and about the interpretation of events, ideas, and experiences that re influenced and shaped by the groups within which they live” (Frost, Moore, Louis, Lundberg & Martin, 1985, p. 17), and “A pattern of shared basic assumptions that the group has learned as it solved problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore to be taught to new members as the correct way to perceive, think and feel in relation to these problems” (Schein, 1992, p. 12).

The organizational culture perspective challenges the traditional structural perspective for analyzing organizations in that 1) organizations are more than structures developed to achieve rational ends and 2) members’ behavior is driven by these factors rather than simply rules or authority and this behavior drives the social life of the organization (Bjerke, 1999; Ott, 1989). Organizational culture is a social construction that is both a product and a process. It’s a product because it is constructed by humans in the form of accumulated wisdom; it is shared with others who learn it, more noticeably

regarding new members of the organization; and it's a process because it gets renewed and recreated (Bjerke, 1999; Bolman & Deal, 2003). The renewal and recreation of culture is part of the natural evolutionary process of the social construction of reality results in humans having a continuing sense of what reality is all about and how they should act upon it (Berger & Luckman, 1965; Bjerke, 1999). Meanings in the form of terms, categories or actions that are accepted by an organization are relevant to a particular place in time, a "snapshot," as this social construction evolves from changes in individuals, the organization itself, its environment, and the advancement of knowledge and technology.

Culture has its strength when it is internalized within the personality of the organization, and this overrides attempts to use it strategically for desired ends. This point is relevant and critical concerning "the culture for evaluation" within an organization where the leadership breeds and infuses a culture of being self critical individually and organizationally. Schein's point about basic assumptions, the things that are engrained in and are a natural part of the organization, is important to understanding how evaluation may become part of an organization's culture. The common phrase "It's just a part of what we do" is a good example of how an organization's stakeholders describe a basic assumption. Bjerke describes this concept as *hidden assumptions* which he defines as "...the fundamental beliefs behind all decisions and actions – that might be nonconscious cornerstones of culture" (Bjerke, p. 34).

In a NHSO we can expect the incorporation of values, focus on the mission, and a commitment to address a particular social problem or human need as likely basic or hidden assumptions. In an operational context, we can also expect fundraising, financial

management and human resources management to be fundamental basic assumptions for a NHSO, as they inherently would perform these functions to survive as open systems and remain competitive in their environment. This raises the questions for 1) how program evaluation can be an integral part of an organization's values, norms and work towards its mission and 2) how it can subsequently become an operational basic assumption at the same level of attention and priority as these other management functions.

The evaluative aspects of an organization can become more significant when value orientations are internalized and role expectations are persistently institutionalized (Parsons in DiMaggio & Powell, 1991). A NHSO has core values emanating from its mission. These values are internalized by the leadership and other stakeholders of the organization. The expected roles of these stakeholders are established through this internalization of these values. If program evaluation becomes part of the internalized values of the organization driven by its mission and part of the expected roles of key stakeholders, it is likely to become engrained in the organization's culture as a basic assumption. Sanders aptly calls this the "mainstreaming" of evaluation and describes it as "...the process of making evaluation an integral part of an organization's everyday operations" and "...part of the organization's work ethic, its culture, and job responsibilities at all levels" (Sanders, 2003, p. 3).

It is helpful to broaden our understanding of culture in order to better understand how program evaluation can be mainstreamed in a NHSO. Culture also consists of "symbolic vehicles of meaning, including beliefs, ritual practices, art forms and ceremonies, as well as informal cultural practices such as language, gossip, stories and

rituals of daily life” (Swidler, 1986). Culture evolves into more of a set of habits, skills and a style of behavior as people adjust their conduct to satisfy their aspirations (Swidler, 1986). The knowledge of their environment becomes important for their understanding of what behavior is acceptable and most effective under what circumstances for helping them meet those aspirations.

Swidler states that the moral work of social movements has been more enduring, in Protestant culture at least, than the ends that the work seeks to accomplish. This poses an interesting question about NHSOs and how they view their work. Do stakeholders such as staff, board and volunteers emphasize the values behind the work, and how it is organized and carried out (the means) more so than the mission and outcomes of the program (desired ends)? If so, can these organizations still achieve these desired ends in spite of such emphasis on the means, and does this emphasis contribute to or detract from the interest in evaluating programs to determine the results of those intended ends?

Swidler’s analysis of organizational culture consists of three steps. First, culture is seen as a “tool kit” in which people draw stories, symbols, etc. to solve problems. Second, the focus is on culture’s causal effect on how people choose their strategies for taking action. Third, it strays from the traditional focus on ends of the action taken and instead focuses on the cultural components that take part in developing those strategies of action (Swidler, 1986). The great diversity, and sometimes contradictions, within cultures results in there being no consistent system or pattern that moves action in a consistent direction. Hence, the “tool kit” Swidler references as the first step in her approach exists as the set of resources that individuals draw upon for constructing strategies of action. People simply may draw differently from their resources, have

different amounts of resources to draw from and decide upon different strategies of action.

Swidler's following statement can be interpreted to help explain the widely discussed emphasis of nonprofit organizations on inputs, activities and outputs (means) in lieu of the intended outcomes of the program (ends):

“If culture influences action through end values, people in changing circumstances should hold on to their preferred ends while altering their strategies for attaining them. But if culture provides the tools with which persons construct lines of action, then styles or strategies of action will be more persistent than the ends people seek to attain. Indeed, people will come to value ends for which their cultural equipment is well suited” (Swidler, 1986, p. 277).

If a NHSO is focused on its end values i.e. mission and program outcomes, it should remain focused on them in spite of a changing environment or circumstances that warrant different approaches (means) to attaining them. If the culture, however, drives the action through ritual, habits, networks, etc. then these strategies of action (means) will attract more focus and priority than the ends. In some cases where this focus is especially strong, the ends may be foregone entirely or may be determined by the organization's cultural capacity to develop such strategies of action intended to attain them. In extreme cases, nonprofit organizations may stray from their mission and their intended outcomes of their programs.

We can glean more insight from Swidler into the possible factors behind the behavior of NHSOs with respect to program evaluation. She states that people are “reluctant to abandon familiar strategies of action for which they have the cultural equipment” (Swidler, 1986, p. 281). If NHSOs do not have the cultural equipment or cultural environment to foster program evaluation, then they are likely to avoid endeavoring in it. Such equipment and environment would comprise the organizational

context for evaluation capacity building as described by Volkov and King that includes “leadership support, stakeholders championing the cause, sufficient organizational demand and interest, tangible incentives for participation, formal training or professional development in evaluation, building of trust, and existence of feedback mechanisms” (Volkov & King, unpublished, 2005).

March describes culture’s influence on behavior that can help us understand why NHSOs endeavoring in program evaluation are most likely responding to an *external pull* for it rather than providing an *internal push*. “Human decision makers routinely ignore their own, fully conscious, preferences in making decisions. They follow rules, traditions, hunches and the advice or action of others” (March, 1978 cited in Swidler, 1986, p. 282). A possible reason behind the resistance or reluctance to engage in program evaluation is when an NHSO might view it as an “organizational ideology for a special cadre within society” (Swidler, p. 284). Quite often leaders of nonprofits have stated they have engaged in program evaluation because their funders have required it. If the call for evaluation is viewed by these leaders as an ideology from the cadre that includes funders who hold large amounts of money and typically do not deliver programs, then that ideology is not likely to be absorbed into the NHSO’s culture and organizational context, or at minimum at the surface level for reasons of appearance or quality signaling. Another example of how this perspective may take hold is when NHSOs, particularly smaller ones, view program evaluation as a luxury for the cadre of larger NHSOs with resources.

There are aspects of an organization that can provide us with the by-products, or evidence, of the organization’s culture. Waterman, Peters and Phillips provided variables

for organizational change in their “7-S Framework” in an attempt to move discussion beyond the influence of structure (Waterman, Peters & Phillips, 1980). These variables represent aspects of a NHSO where we can look for cultural products that support the institutionalization of program evaluation, as shown below in Table 2:

Table 2 – Cultural Products for Program Evaluation by Organizational Variable

Organizational Variables*	Cultural Products
Structure	<ul style="list-style-type: none"> • Program evaluation is a permanent function as opposed to being performed ad hoc. • A staff person is dedicated to the function of program evaluation, directly performing it or working with external evaluators. <ul style="list-style-type: none"> ○ Staff person reports to the ED or senior management.
Strategy	<ul style="list-style-type: none"> • Program evaluation drives, or is a prominent part of, the organization's strategic planning process. • Evaluation plans are used for all program evaluations conducted. • An evaluation capacity building plan is used to develop initial or expand existing program evaluation capacity. • Program evaluation efforts are part of the organization's strategic response to changes in the environment regarding – <ul style="list-style-type: none"> ○ demands from external stakeholders ○ program consumer needs ○ program design, logistics and delivery ○ improving performance to remain competitive
Systems	<ul style="list-style-type: none"> • Program evaluation is specifically referenced in the organization's budget as a line item or function. • Policies and procedures specific to program evaluation are established and followed. • Data collection systems are driven by the requirements of the program and evolve as programs, external data (reporting) demands and evaluation needs change. • Communication systems such as feedback loops exist to ensure communication flows fluidly and consistently among stakeholders.
Style	<ul style="list-style-type: none"> • The ED personally supports program evaluation through explicit actions, as well as rhetoric and symbols. • The organization's leadership (Board, ED and senior management) regularly reinforce the importance of program evaluation to the organization and the consumers it serves. • The leadership and management styles contribute to developing a culture for continuous improvement supportive of program evaluation.
Staff	<ul style="list-style-type: none"> • Program evaluation influences hiring choices with regard to program personnel having an evaluation background, or having the skills to engage in the evaluation process. • Program evaluation is used to boost staff morale by giving new or enhanced meaning of their work. • Program evaluation is linked to staff performance when appropriate.

Skills	<ul style="list-style-type: none"> • Program evaluation serves as the impetus to provide professional development to enhance the skills of staff. • Program evaluation results are viewed as a representation of what the organization does best.
Super-ordinate goals**	<ul style="list-style-type: none"> • Program evaluation is part of the organization's values, future direction, and represents how the organization wishes to be viewed. • Program evaluation has special meaning and serves as a rallying point for stakeholders through its connection to the organization's mission.

* From Waterman, Jr., R.H., Peters, T.J. & Phillips, J.R. (1980). Structure is not Organization. *Business Horizons* 23(3), pp. 14-25.

**Defined by Waterman, Peters and Phillips as “guiding concepts – a set of values and aspirations, often unwritten that goes beyond the conventional formal statement of corporate objectives” p. 25.

Table 2 serves to demonstrate how organizational culture plays an important role in ECB. When program evaluation is part of an organization's *structure*, in terms of it being a permanent, dedicated function appropriately placed in the organization's hierarchy, it likely is evidence that evaluation capacity has been built and will be continue to be built in the organization. *Strategy* plays an import role in ECB in three areas. First, at the organizational level, if program evaluation is part of the organization's strategic plan it is part of the organization's goals and objectives. While implementation of the plan is no guarantee, program evaluation is integrated in the desired future state of the organization and therefore stands a better chance of having capacity built for it. Second, program evaluations need to be planned for several reasons. Evaluability assessments at the organizational and program levels help determine what is necessary to efficiently and effectively evaluate programs. This up front planning helps and organization determine the feasibility of evaluating programs and provides a guide for acquiring the necessary elements to conduct the evaluations. Third, the process of ECB should have its own plan that is monitored, updated and communicated to stakeholders.

An ECB plan helps an organization stay focused on building capacity, ensure stakeholder accountability, and incorporate the process into organizational learning.

Successful ECB requires specific *systems* that program evaluation is a part of and systems it requires to be functional. Program evaluation needs to be budgeted, part of the organization's policies and procedures, and requires appropriate data collection and communication systems driven by what is necessary to evaluate specific programs. The leadership and management *style* of the organization can ultimately determine the success of that organization's ECB. Prioritization, stakeholder buy-in, and subsequent marshalling of resources for program evaluation are just some of the important cultural drivers for ECB that leadership directly impacts which will be discussed further in the next section. The role of *staff*, particularly program staff, is critical for ECB. Without the support, cooperation and participation of the program staff in a NHSO, the evaluator is highly unlikely to be able to perform the evaluation. Conversely, program evaluation can be effectively used as a tool to boost the morale of staff by establishing, refreshing, or renewing their perspective for their work, the work of the organization and their role in impacting their consumers. Staff has to have certain *skills* to be able to effectively work with evaluators, so organizations that recognize this and promote opportunities for professional development contribute to the success of ECB and foster organizational learning.

Lastly, Waterman, Peters and Phillips discuss what they call *super-ordinate goals* as the guiding values and concepts for an organization (1980, p. 25). NHSOs tend to be driven by their missions and a set of values based on the service they provide and who they provide it to. Cultural products indicating program evaluation has become part of

the organization's super-ordinate goals can be when it becomes a basic assumption or inherently part of what the organization does; a connector, driver and rallying point for the organization's work towards its mission; an explicit demonstration of the organization's philosophy for program evaluation; and the incorporation of program evaluation into the organization's values and efforts to achieve its desired future state.

These variables and corresponding cultural products are not exhaustive for representing cultural evidence of the institutionalization of program evaluation. The following sections will attempt to further examine organizational culture's role in evaluation capacity through organizational learning, stakeholder involvement and leadership.

Leadership and Organizational Culture

Leadership has been defined in many ways throughout the vast literature on the subject. For the purpose of this research Northouse's definition will be used, "Leadership is a process whereby an individual influences a group of individuals to achieve a common goal" (2004, p. 3). This definition provides features of leadership congruent with those of program evaluation in NHSOs. Both are a process, involve influence, occur within a group context and involve the attainment of goals. These parallels provide some clues as to why the role of the leader and their influence might be important for building evaluation capacity in their organization.

The culture of an organization is demonstrated in many ways, but one of the more prominent is the management style of the organization. A nonprofit organization's ability to successfully attain its goals and objectives while working toward satisfying its mission largely relies on the leader's (in this context, the ED) ability to effectively work

with key stakeholders such as the board, staff and operational volunteers. This involves the ED creating purpose and commitment for stakeholders (most notably board, staff, and volunteers) which is arguably the leader's most influential impact on organizational culture (Bjerke, 1999). The organization's mission serves as the leader's driver and rallying cry for such purpose and commitment; the guide for strategy, performance and organizational change; and the representation of the social goods a NHSO provides for society.

The leader of an organization can have profound influence on their organization's culture in many ways. The culture of a nonprofit organization is largely built upon the ED's values, activities and tasks which are inculcated to staff and other stakeholders (Hay, 1990). The leader's use of language is a key driver of the acculturation process. The way values and rules of behavior are communicated, including nonverbal communication, shapes the values and norms of the organization's culture (Bjerke, 1999). A leader can have formal influence based on their position and authority (Ivancevich, Szilagyi, Jr. & Wallace, Jr., 1977) and set the priorities for the organization. He or she may also have informal influence based on their expertise or special skills that are important for the organization, such as an ED with expertise in fundraising or financial management.

The issue of accountability has garnered much attention recently in the nonprofit sector. While it typically is linked to more top-down approaches to management and leadership within the context of meeting goals or objectives, it also influences the culture of an organization. If a leader communicates the rules of behavior and the priorities for an organization, this communication has no meaning for the organization's members if

there is no accountability for following those rules or satisfying those priorities. Without the proper accountability framework, a leader's voice to the organization becomes rhetoric without any foundation for meaningful action. The leader must provide the organization's members with the level of support and working environment that enables them to deliver on their area of accountability. This involves leveraging the creative capabilities of the organization's members, engaging members through the psychological contract of their expectations combined with those of the organization, aligning members' thoughts decisions and actions with their goals and the roles involved in achieving them, and developing members to help them realize their potential through mentoring and coaching (Kraines, 2001).

Leaders in and of themselves are symbols, and their patterns of behavior and leadership can be an artifacts for the organization that communicate information about the organization's values, guiding beliefs and ways of doing things (Davis, 1984; Ott, 1989). Congruence must exist between the cultural values and operating norms for an organization to be successful (Anthes, 1987). Impacting an organization's culture will likely influence the organization's strategic direction, and the ability to achieve its goals and objectives (Davis, 1984; Hay, 1990). Effective leadership and a supportive culture are typical characteristics of high performing organizations and specific leader practices include permeating a strong customer (consumer) orientation throughout the organization, demonstrating a strong commitment to quality, involving and empowering staff by ensuring they participate in decisions and that their feedback is used, and ensuring employees have the proper training to perform their jobs well (Wiley & Brooks, 2000).

In order for leaders to be enablers of a successful execution of strategy, they must align the values and culture that support that strategy (Kaplan & Norton, 2006). More specifically, an organizational culture that promotes an environment for employee satisfaction is more likely to enhance productivity and organizational effectiveness (Ostroff & Schmitt, 1993). The hierarchical position of EDs aids their ability to shape and change the culture through serving as role models, espousing their values and interacting with their stakeholders while ultimately facilitating or limiting those stakeholders' actions (Cooke & Szumal, 2000). A leader can be the driver for cultural change within an organization, but only if he or she can handle personal and/or professional discomfort that typically comes with the reflection and introspection necessary for effective change (Block, 2004).

Transformational leaders are those who engage and connect with stakeholders to attend to their needs, their motives, help them reach their full potential and together develop meaning and purpose for the organization and its work (Bjerke, 1999; Northouse, 2004). An effective transformational leader can help stakeholders feel they are empowered. Some indicators for this include stakeholders feeling they are significant, they contribute to the organization's work, they feel part of a social community and they generally are more enthusiastic about their work (Bennis, 1989). Transformational leaders are differentiated from transactional leaders who focus more on their exchanges with stakeholders such as providing incentive and reward systems, decreeing new policies or procedures, or altering agendas. The transformational leader seeks to transcend values for the purpose of uniting stakeholders for organizational change. If we view program evaluation can be a vehicle for organizational change; providing new

meaning for the work of stakeholders; creating an environment and culture conducive to organizational learning and continuous improvement; and assessing the organization's work towards its mission, then we can hypothesize that transformational leaders will be more effective at building capacity for it, mainstreaming it and institutionalizing it.

There remains continuing debate for whether organizational culture can be managed, and constraints for success in this process include resistance from stakeholders, organizational life cycles, poor communication, lack of stakeholder development, subjective interpretations of issues and bad timing (Nord, 1985). There is also debate over whether the leader is managing the culture or the culture of the environment is managing the leader. However, what remains constant is that a leader of an organization can have great influence on their organization's culture and be the catalyst for cultural change. If a leader develops and fosters a culture that is positive, constructive, values knowledge, skills, innovation and continuous improvement it is likely to motivate stakeholders and contribute to organizational learning. More specifically, it is also likely to value, encourage and reward continuous professional development (CPD) for program staff (Thompson, 2000). Conversely, if a leader fosters a culture of negativity and cynicism, it is likely to serve as a barrier to successful outcomes for consumers and endeavors in processes requiring critical thought such as reflective practice and program evaluation. The importance of how a leader influences their organization's culture can be observed in how it impacts the organization's vision, strategy, direction, operations and performance. This influence might also contribute to ECB and building a culture for evaluation. The ED's specific role in ECB will be discussed in the Roles of Stakeholders section.

Managing the Socio-Political Environment

Subjectivity has a role in research (and in program evaluation), and those engaged in either simply may not be in agreement on subject matters, results or predictability due to the fact that value judgments are part of the equation. Schumpeter cautions researchers about positing ultimate ends and “not preventing (those) ends and sympathies from influencing the results” (Schumpeter, 1991, p. 318). These ultimate ends can be translated into the long term outcomes that NHSO program espouse to attain for their consumers. Subjectivity is inevitable in program evaluation, as quite often with so many variables involved in what is being evaluated, especially in the complex nature of NHSOs, “evaluation results are interpreted subjectively, and different people can interpret the data in many ways” (Murray, 2005, p. 352). Weber reminds us however that just because value judgments are subjective, they should not be removed from scientific debate. Therefore, the problem, as he puts it, is not the value judgments themselves but “the meaning and purpose of the scientific criticism of ideals and value judgments” (Weber, 1949, p. 52).

Subjective interpretation is driven, in part, by our individual social construction of reality. What is viewed as reality or knowledge is driven by its social relativity and the processes included in the “sociology of knowledge” (Berger & Luckman, 1967, p. 3). For example, what may be real to one person in their culture may not be real to a person from another culture. The knowledge from one profession may differ from that of another profession based on the social context of each. Beyond these differences, the sociology of knowledge must also deal with how knowledge becomes established as reality, and we’re reminded that “all knowledge is developed, transmitted and maintained

in social situations” (Berger & Luckman, 1967, p. 3). Therefore, the sociology of knowledge is concerned with how knowledge is constructed into reality.

Everyday life is essentially analyzed and subjectively interpreted by people. This subjective interpretation is played out in people’s thoughts and actions that help them maintain their sense of reality. Even as individuals, our worlds consist of multiple realities as our attentiveness moves to different levels within our consciousness. Everyday life is therefore experienced in “differing degrees of closeness and remoteness, both spatially and temporally” that a world of inter-subjectivity that we share with each other (Berger & Luckman, 1967, p. 22). People may not view the world in the same way as they have different or overlapping senses of their “here and now.”

The accumulation of our experiences, specifically what is retained, is referred to as “the social stock of knowledge” (Berger & Luckman, 1967, p. 41). We continually contribute to and draw from this social stock of knowledge in everyday life. It reminds us of our current situations and our limitations. For example, we can be aware of who is employed and who is not. This point has clear implications for NHSOs and the consumers they serve. Berger and Luckman amplify this point by stating that “participation in the social stock of knowledge thus permits the location of individuals in society and the handling of them in the appropriate manner” (Berger & Lukman, 1967, p. 42). NHSOs typically provide a specific service to consumers in their “location” such as the homeless, the hungry, those involved in substance abuse, those who have a disability, etc.

Social stocks of knowledge help us, and presumably NHSOs, know what to do in certain situations. Our direction and actions are driven by our own logic which is based

on what we know. It is important to note that as the stock of knowledge changes, seemingly the theory of change (or logic) for a NHSO program also would change based on testing or evaluating the program. This can be largely driven by EDs when they enable imperfect information about their programs to be used as the constructive means for program improvement.

The social distribution of knowledge is not even and each of us possesses different amounts and types of knowledge. Accordingly, we choose who to share our knowledge with and at times to not share it with anyone. As a result of the deficiencies of knowledge we don't possess, we need the knowledge and advice of experts. Berger and Luckman tells us that in our everyday life, "we know what (information) we can hide from whom, whom we can turn to for information on what we don't know and generally what types of individuals may be expected to have which types of information" (Berger & Luckman, 1967, p. 43). This has implications for whether an ED decides to hire an evaluator to compensate for lacking the internal organizational expertise necessary to properly evaluate its program.

Evaluation involves making a value judgment for a program and is therefore inherently a political process (Fitzpatrick, Sanders & Worthen, 2004; Palumbo, 1987; Weiss, 1998). Program evaluation operates in a political environment where it is one factor for making decisions (Chen, 2005). In the broad context of human services, the determination of what problems society feels are important enough to be addressed is a political process driven by policy, interest groups, and the support of political leaders (Monette, Sullivan & DeJong, 2002). The political environment for program evaluation focuses on programs that were developed through political decisions (Weiss, 1987).

Nonprofit programs are typically developed because a person or group of persons determined there was an unmet need in society.

Program evaluation results enter a political decision making process with regard to how to use them, how to present them and the priority they carry when competing with other demands of time and attention from decision makers (Fitzpatrick, Sanders & Worthen, 2004; Weiss, 1987). Evaluating programs for public relations purposes or to fulfill grant requirements are some examples of political reasons why programs are evaluated (Weiss, 1998). Other examples of political issues stakeholders should be aware of include the priorities and distribution of resources, evaluators being co-opted by the organization contracting with them, and actions taken to subvert the evaluation (Fitzpatrick, Sanders & Worthen, 2004).

Program evaluation operates within the sociopolitical context in which human services are practiced that includes agendas, interests, competition, sponsorship and the control of resources (Tripodi, 1983). This context includes internal stakeholders such as the ED, Board of Directors, staff, volunteers and clients as well as external stakeholders such as government, funders, accreditation organizations, the community, businesses, the media, and others. Stakeholders are advised to recognize this environment and make attempts to effectively manage it when attempting to build capacity for program evaluation and evaluate programs. Fitzpatrick, Sanders & Worthen remind us that “Many a good evaluation, unimpeachable in all technical details, has failed because of interpersonal insensitivity, ethical compromises, or political naiveté” (2004, p. 411). The program evaluation standard for political viability states that “The evaluation should be planned and conducted with anticipation of the different positions of various interest

groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted” (Joint Committee on Standards for Educational Evaluation, 1994).

Another aspect of this environment is determining whose interests are served by having the programs evaluated, which is inherently a component of any process seeking to measure effectiveness (Pfeffer & Salancik, 2003). Consumers of evaluation are defined as people or organizations that call for it and use its results in some way (Tripodi, 1983). These stakeholders may include the ED, the Board of Directors, the funding organization, the clients, the program staff, the accreditation organization, and a government agency. These interests can compete and conflict with each other making the management of the political environment extremely difficult. Abma suggests three important political challenges that include asymmetric power relationships, sensitive issues with consumers, and strategic behavior’s potential for smothering open or appreciative inquiry (2006).

There are various instances where stakeholders will play a role that accentuates the political environment for program evaluation or help manage it to reduce the political impediments to the evaluation capacity building process. For example, nonprofit organization staff, which may include the ED, may attempt to have evaluation results be presented so their organization is represented favorably (Mohan & Sullivan, 2006). Program staff may have a false vision of program evaluation as a scientific process removed from politics (Reamer, 1998) and can be disappointed and discouraged when engaged in the process. Conversely, they may feel apprehension towards it, especially in

cases where the programs are being formally evaluated for the first time and when they have delivered the program for a substantial length of time. This scenario is not exclusive to relationships with external evaluators, as internal evaluators may also be treated as “outsiders” trying to tell the people who deliver the program and work intimately with the consumers how effective “their” program is doing.

The relationship between the program staff and the ED can play an important part in this political arena when one or the other initiates program evaluation. If the ED and the program staff have a good relationship with each other, then the process is more likely to progress smoothly with harmonious support. When the relationship is not cooperative, regardless of who calls for it, the process is less likely to be efficient or effective. This relationship is discussed in more detail in the stakeholder roles sections for evaluators and program staff.

Evaluators face political challenges in dealing with a myriad of stakeholders. For example an ED, who also might be signing the checks to pay the evaluator, may request that what they perceive as undesirable information or information that reflects bad on the organization be omitted from the final report. Many NHSOs receive a substantial portion of their overall funding from government agencies, and as administrations and priorities change so do the funding interests. As a result, the NHSOs addressing the social problems that the current administration has deemed important are most likely to receive funding and support for program evaluation. The process of receiving private funds from foundations is also not removed from politics, as their program officers, EDs and board members may all have a hand in determining the funding priorities from the level of a social problem down the level of what NHSOs and what programs they will support.

An ED who desires to effectively build capacity for program evaluation in their organization must be aware of and understand the socio-political environment in which evaluation exists. Researchers make choices that are subject to “the horizon of the analyst ...the information and mental equipment at his command” (Schumpeter, 1991, p. 317). This idea certainly applies to the area of capacity for program evaluation from the evaluator’s and the NHSO’s stakeholders working directly with the evaluation process. All involved come with their own “horizon,” “information,” and “mental equipment.” A common thread through the evaluation capacity building (ECB) literature that emanates from the evaluation field is the issue of skill sets of the evaluator and the program’s stakeholders involved in the evaluation process. Examples from the Evaluation Capacity Building Checklist include:

“Establishing a capable evaluation oversight group (composed of members of the staff, board of directors, and community) to initiate, continually evaluate, and advance evaluation processes in the organization” and “building up organizational and individual ability and readiness to implement evaluation activities” (Volkov & King, 2005).

EDs may take several actions in an attempt to manage the political environment for program evaluation. It is important for the ED to understand that this political environment is not strictly bounded, but it is rather permeable and susceptible to outside forces (Mohan & Sullivan, 2006). The ED must acquire a thorough understanding of this political environment by identifying the stakeholders, their interests, and their interactions. In addition to internal stakeholders, these may also include program sponsors or funders and any relevant legislation or policy linked to the program(s) (Scriven, 2007). Once the key stakeholders have been identified, the ED can begin to

strategize for how to interact with them in an attempt to manage the political environment towards maximizing evaluation capacity.

A leader who embodies the ideals in their organization's mission through their actions and ensures the mission is part of their evaluation criteria (Brinckerhoff, 2000) can explicitly link the effort to evaluate programs with working towards satisfying that mission. Making this connection part of the organization's culture can ensure program evaluation is a priority for stakeholders such as staff, the board and other volunteers. The mission can be a powerful rallying cry for building evaluation capacity, as it is quite often what attracts stakeholders to be employed or volunteer at nonprofit organizations. Utilizing the organization's mission can help diffuse political obstacles such as personal agendas, fear of evaluation, and fear of information perceived as a negative reflection on the programs, the organization and/or the individual stakeholder's work. It is the ED's role, as the top hierarchical staff person, to develop a risk-friendly environment that is without fear of negative consequences, to nurture a culture of trust, and to sincerely value honest and open feedback as a means for effectiveness (Gray & Stockdill, 1995).

An ED can help develop a culture for continuous improvement by ensuring that stakeholders understand that there is no such thing as "negative" information about the organization's programs and that all information represents opportunities for learning and improvement. This effort can be the beginning of an ED's process of utilizing program evaluation as the means for organizational learning (Gray & Stockdill, 1995; Presskill & Catsambas, 2006). Becoming a learning organization, according to Green, is one of the ten most important things nonprofits must do to "...survive, adapt and thrive" in this new century (2004, p. 20). Evaluation becomes a part of organizational learning when it

becomes institutionalized as part of the sources of an organization's information, power structure, processes and systems that influence decision-making and action (Boyle & Lemaire, 1999). The ED through managing access to and dissemination of information is in an optimal position to influence the institutionalization process. The ED, based on his or her responsibilities, hierarchical position, connection to the board of directors and influence that can reach throughout their organization, is also in an optimal position to impact the building of capacity to evaluate their organization's programs.

A reoccurring theme in the literature for strategies attempting to contend, if not manage, the socio-political environment for program evaluation is how stakeholders are involved in the process. This is because political obstacles to effective and efficient program evaluation are driven and magnified by uncertainty and confusion (Chen, 2005). Examples of stakeholder involvement that help reduce these factors and reap other benefits are discussed in the section *Stakeholder Roles*.

Meeting the Challenges for Program Evaluation

Nonprofit organizations that desire to engage in program evaluation face challenges such as the lack of skills or expertise, funding, or other resources (Dym & Hutson, 2005; Paddock, 2001). Some organizations depending on their niche, service delivery and external stakeholders face the demand for more sophisticated methods of evaluation (Kearns, 2001). Other challenges include political tension between evaluation stakeholders (Murray, 2005; Oster, 1995; Paddock, 2001; Thomas, 2005) and potential incongruence on what comprises effectiveness (Herman & Renz, 2002). Lastly, organizations may simply feel overwhelmed by the overall complexity of the program evaluation process (Oster, 1995). The ED's role in the overall process of program

evaluation is receiving increasing attention and some researchers and scholars indicate that the ED's support of the evaluation process is critical to its success (Campbell, 2002; Mesch & McClelland, 2006; Oster, 1995; Thomas, 2005; United Way of America, 1996). Specific examples of what EDs can do to ensure the process is successful include developing goals with program managers that set the context for program evaluation (Oster, 1995), initiating the evaluation process (Thomas, 2005), and forming evaluation work groups (United Way of America, 1996; Werther, Jr. & Berman, 2001).

Overall NHSOs face the complex issues of evaluation capacity such as context, resources structure, balancing external demands with internal support, influencing organizational culture, and managing the socio-political environment. However, there is also the challenge of realizing that evaluation is not a remedy for all of society's problems. Not all evaluations of programs will successfully lead to the improvement of those programs due to a myriad of reasons. Program evaluation comes with its limitations, and they can be magnified by one or more key stakeholders in the process. One example is when evaluators promise an organization they can acquire specific desired information about their program or evaluation results that cannot be attained (Fitzpatrick, Sanders & Worthen, 2004). This can not only result in an unsuccessful evaluation but also yield program information that does not meet the evaluation standards for *accuracy* or *utility*. Such consequences can have devastating effects on the evaluation capacity building process and the stakeholders involved, especially for organizations venturing to evaluate their programs for the first time. A considerable amount of effort may be extended to build the culture and support for evaluation only to have it sabotaged by a bad first experience, resulting in the possibility that the organization may not engage

in evaluating its programs again at all or at the necessary level of rigor and comprehensiveness. The evaluation profession also does not benefit when organizations have not had good experiences and convey them to their peers.

Conversely, the NHSO can magnify program evaluation's limitations by not having realistic expectations for it. First, it can expect it to solve all of the organization's problems, especially when it views it primarily as a process to meet the external demands of stakeholders. Such a perspective can result from 1) the NHSO stakeholders not making an effort to adequately understand evaluation and what their role is in the process and/or 2) use program evaluation to respond to unrealistic demands from external stakeholders such as measure outcomes that do not match well with their programs or mission.

Second, it may misunderstand the evaluation results, especially for quantitative data, to comprise the final answers to the program's theory. In most cases, results, especially numbers, raise more questions than produce answers. The purpose of program evaluation is to identify the program's strengths and weaknesses and provide helpful information so stakeholders can learn more about their programs (Fitzpatrick, Sanders & Worthen, 2004). Program evaluation within the context of continuous improvement is not a one-time process, as NHSO stakeholders are responsible for using the information to make the necessary changes to the program with the ultimate goal of better serving their consumers.

Third, a NHSO can decide to develop programs not related to their mission because that particular program area is what the current funding environment is supporting. This misalignment of programs and mission, in addition to representing

mission creep, can implicate the evaluation process in several ways. It can be difficult to adequately evaluate programs that are not part of the normal business of the organization because the program theory, infrastructure or delivery systems have not been adequately developed. The organization's culture has been built upon its original mission that its staff and volunteers have bought into and supported. This level of enthusiasm and engagement might not be present in a program misaligned with the organization's mission. Without such support from these stakeholders the evaluation process will prove difficult.

Courses of Action for Program Evaluation

The environment in which nonprofit organizations exist is constantly changing, and EDs face the challenge of conforming to environmental constraints while attempting to develop a more favorable environment for their organization (Pfeffer & Salancik, 2003). This leadership and management challenge suggests that leaders of organizations must constantly measure their organization's performance (Nanus & Dobbs, 1999). EDs of nonprofit organizations continue to be challenged by an environment that is increasing in competition for resources and the call for accountability. Program evaluation is playing an increasing role in these shifts, as program effectiveness has been factored into funding decisions and is balancing fiscal responsibility for overall organizational accountability. The results have been an increasing call for the demonstration of program and organizational effectiveness (Brinckerhoff, 2000; Carlson & Donohoe, 2003; Gray & Stockdill, 1995; Mesch & McClelland, 2006; Werther, Jr. & Berman, 2001). These external demands comprise only one aspect of what might be driving an ED's course of action with regard to program evaluation.

Studies on human conduct are typically framed with means and ends (Weber, 1949). We either desire something as an end or as the means for achieving something else that is also desired. Science does allow us the ability to determine the appropriateness of certain means for achieving certain ends. We can also determine the consequences of enacting certain means and their cost for achieving desired ends. However, science is not equipped to analyze this framework for decision making, as that is something we as participants in the process must do. We are the ones making the rational choices based on what can be a myriad or small set of criteria. As Weber states, “he weighs and chooses from among the values involved according to his own conscience and his personal view of the world” (Weber, 1949, p. 53). Science’s role in this process is to help the social actor realize that there are choices, those choices have consequences and that making the choices is his/her responsibility.

Value judgments come into play when social actors set goals with desired ends and test the various means used to attain them. People can even rank choices of means in order of cost, and in doing so they develop standards. Weber tells us that not science, but only the person’s will and conscience will determine whether those standards will be adhered to. Leaders of NISOs must decide if endeavoring in evaluating their programs is worth the costs, if so to what level of comprehensiveness and ultimately whether their organization will maintain such standards of comprehensiveness necessary to validly measure the outcomes of their programs. Weber states that “an empirical science cannot tell anyone what he should do – but rather what he can do – and under certain circumstances – what he wishes to do” (Weber, 1949, p. 54).

Evaluating NHSO programs involves distinguishing between value judgments and empirical knowledge. The objectivity of what is considered truth is the driver of this dilemma. Weber mentions familiar potential complications for program evaluation such as “conflict about methods ...the incessant shift of viewpoints and the continuous redefinition of concepts” (Weber, 1949, p. 63). There is also the common misunderstanding of evaluation results containing all the answers for a program and for it being an exact science, typically neither of which it is. Weber reminds us that we are not concerned with “laws” but with ‘adequate causal relationships expressed in rules and with the application of the category of objective possibility” (Weber, 1949, p. 80). Program evaluation results typically tell us what happened, but only with further investigation can tell us how things happened or why they happened. Issues of subjectivity, lack of objectivity, lack of agreement on methods, intended outcomes and other parameters for program evaluation all can contribute to discouraging leaders of NHSOs endeavoring in program evaluation.

There are four primary courses of action EDs can choose with regard to meeting the challenge for demonstrating the effectiveness of their organization’s programs. First, they may choose to do nothing for several reasons. If their organization is fiscally healthy, and they view program evaluation solely as a means to satisfy an external stakeholder and acquire financial resources, they may not see the need to demonstrate program effectiveness. This choice may also be predicated on determining that the cost of evaluation process is greater than the financial resources acquired. An ED might desire to have their programs evaluated but may feel forced to do nothing because their organization does not have the capacity.

Second, they may choose to engage in evaluating their organization's programs for the sole purpose of meeting the demands of external stakeholders such as funding organizations, government agencies and/or accreditation organizations. This is typically done for the purposes of acquiring resources or satisfying regulatory, accreditation or contractual requirements. This can be characterized as an external pull. The ED has chosen to have their organization respond to the demands of external stakeholders and therefore allow those demands to drive the evaluation process. Resource dependency is a powerful influence on nonprofit organizations since they are open systems that rely on resources from the external environment. However, an organization that continuously engages in this external pull can run the risk of losing autonomy and drifting from their mission. Cooke and Szumal characterize this as a "culture disconnect" where external demands and the need for resources "...are more influential in shaping systems and related antecedents than are the espoused values of members or the organization's mission or philosophy" (2000, p. 159).

Third, the ED may decide to develop an intrinsically motivated effort for program evaluation within their organization focused on organizational learning, a culture of continuous improvement, and driven by the organization's mission. This can be characterized as an internal push. EDs typically already are engaged in the first step toward building intrinsic motivation among their internal stakeholders by communicating their organization's vision, mission and strategy (Kaplan & Norton, 2006). The key to embedding program evaluation in this effort is for the ED to communicate that it can be an effective tool used for determining the organization's progress towards accomplishing its vision, mission and strategy.

Lastly, the ED may choose to integrate the demands of external stakeholders from the external pull into the efforts of the internal push in an attempt to initiate, nurture, manage and balance both. The ED that chooses this path has to be effective in balancing the importance of satisfying external stakeholders and acquiring resources with retaining autonomy and the direction of the organization's mission. Balancing the external and internal organizational contexts is important for effective building of evaluation capacity. There must be the inclusion of external demands and support from the organization's leadership (Volkov & King, 2005). Inclusion of the internal push also enables an organization to work towards mainstreaming evaluation, "...the process of making evaluation an integral part of an organization's everyday operations" and "...part of the organization's work ethic, its culture, and job responsibilities at all levels" (Sanders, 2003, p. 3). Evaluation then can become a "basic assumption" for solving problems "...of external adaptation and internal integration" (Schein, 1992, p.12).

Organizational Learning

Organizational learning involves learning from successes and failures, and changing behavior due to encountering situations, usually emanating from its environment. It involves individuals learning new ways to achieve their goals and sometimes involves individuals learning to change the goals (Denhardt, Denhardt & Aristigueta, 2002). Organizational learning occurs when "...individuals within an organization experience a problematic situation and inquire into it on the organization's behalf" (Argyris & Schon, 1996, p. 16). Learning resulting from this inquiry "...must become embedded in the images of the organization held in its members' minds and/or in the epistemological artifacts embedded in the organizational environment" (Argyris &

Schon, 1996, p. 16). While organizations are collectives of individuals, we cannot assume that individual learning permeates and takes hold at the organizational level. Individual learning does not guarantee organizational learning, but it is a prerequisite for it (Senge, 1990). Organizations need to be environments for storing knowledge that goes beyond the individual level in files, within policies, regulations, strategic plans, documented decisions, etc. (Argyris & Schon, 1996). Argyris and Schon also recommend that organizations “directly represent knowledge in the sense that they embody strategies for performing complex tasks that have been performed in other ways (1996, p. 13). In a highly competitive and rapidly changing environment organizational learning must occur quickly and be sufficient enough to match or be greater than the change outside the organization (Schwandt & Marquardt, 2000) and in a way that they can sustain their mission and purpose (Dym & Hutson, 2005).

Leaders are challenged with ensuring that individual members learn, grow and adapt in an environment that encourages and rewards them (Ott, 1989). They should also ensure that they have access to information, a feedback loop exists and that the learning process contributes to improved performance (Mesch & McClelland, 2006). Information must flow in both hierarchical directions for a nonprofit to be a learning organization (Drucker, 1990). Organizational learning is a socialization process driven by the structures, systems and culture. Without encouraging, fostering, nurturing and rewarding an organizational culture for learning that allows for failure and taking risks, an ED cannot successfully lead a learning organization even with the proper structures and systems in place. EDs who wish to lead a learning organization should focus on building an infrastructure for learning that includes recording the organization’s history, reflecting

on successes, failures and innovation, and incorporating learning into the organization's planning process (Senge, 1996). A culture for organizational learning also involves internal and external collaborative efforts with other stakeholders and organizations, and receptiveness to taking calculated risks (Block, 2004), all staples of managing and leading a nonprofit organization. Organizational learning is at the heart of organizational culture's impact on evaluation capacity building. An ED is in the position to ensure that stakeholders are involved in the evaluation process and have an opportunity for meaningful reflection that leads to learning. Patton reminds us that "...participation and collaboration can lead to ongoing, longer-term commitment to using evaluation logic and building a culture of learning in a program or organization" (1997, p. 100).

One particular aspect of organizational learning, "inference of causal connections between actions and outcomes and their implications for future action" (Argyris & Schon, 1996, p. 17) helps describe the process of program evaluation. It appears that program evaluation and organizational learning are inextricably linked. Program evaluation can be an example of and a tool to foster organizational learning, while a culture within an organization that is driven by organizational learning is also one that is conducive for program evaluation. Some suggest that evaluative inquiry, "an ongoing process for investigating and understanding critical organizational issues" (Presskill & Torres, 1999, p. 1), is the means for fostering organizational learning. Evaluative inquiry positions the evaluator as more of a facilitator of learning where he or she works with stakeholders to ensure participation, accountability and a culture that supports continuous learning. This culture enables the evaluator and participating stakeholders to effectively determine a program's strengths and weaknesses through collaborative efforts (Presskill

& Torres, 1999). Preskill and Torres point out that developmental, collaborative, participatory and empowerment evaluation approaches all place the evaluator as a “facilitator of learning” and “emphasize learning as an outcome of the process” (1999, p.183).

Organizational learning also is important to human services as a by-product of the individual learning of program staff (professionals). We know organizational learning inherently requires that the individuals in the organization learn and grow as people and as professional employees. When individual human service professionals engage in research-minded, informed, reflective or evidence-based practice they are taking responsibility for their own learning. Thompson offers several other factors for theory and practice in human services that can contribute to the process of organizational learning in NHSOs. Human service professions evolve by expanding and revising their respective bodies of knowledge, as for example in the case of recommended or “best” practices. Professional practice requires education in addition to training so that practitioners can move beyond updating their skills to acquire critical analysis skills and undertake self-directed learning. A NHSO’s philosophy for developing its staff should be driven by a focus on knowledge, skills and values. Opportunities for a practitioner’s experiences to be targeted to their specific individual learning needs should be exploited to add to promote their enthusiasm and reward them for expanding their will to learn. Lastly, it is important for the NHSO to promote individual attitudes and provide organizational support for continuous professional development (Thompson, 2000, pp. 140-144).

The responsibility for a NHSO becoming a learning organization resides primarily with the organizations internal stakeholders such as the ED, board and staff. Evaluators can help facilitate organizational learning in NHSOs through taking a learning approach to evaluation where these internal stakeholders participate in decision-making, interpreting results and determining next steps. This learning approach involves focusing on important issues and concerns; dialog and reflection about improvement; being able to cope with difficult circumstances for the program and potential “negative” information about the program’s effectiveness; and insightful assessments and discussions about the past, present and future (Torres & Preskill, 2001). These steps also help contribute to a NHSO’s culture for continuous improvement and motivation for program evaluation. Improved program staff practice can lead to an improved intervention, program and consumer outcomes as well as improved overall organizational performance.

Stakeholder Roles

Stakeholder involvement has received increasing attention for its importance in the evaluation and evaluation capacity building processes. The Centers for Disease Control and Prevention references *engaging stakeholders* as the first step in their Framework for Program Evaluation in Public Health to ensure adequate participation, that people’s voices are heard, that their perspectives are understood and to gain their support for the process and for the use of evaluation results (CDC, 1997). Another important reason often cited in support of stakeholder involvement is to ensure their values are incorporated into the process, as they likely to be intimately connected to what would comprise program improvement (CDC, 1997; Mark, Henry & Julnes, 2000). This can be especially important for NHSOs where values are the basis for their organizations,

their programs and the human service professions. The CDC describes stakeholders as, “the persons or organizations having an investment in what will be learned from an evaluation and what will be done with the knowledge” and offers the examples of:

“sponsors, collaborators, coalition partners, funding officials, administrators, managers, and staff, clients, family members, neighborhood organizations, academic institutions, elected officials, advocacy groups, professional associations, skeptics, opponents, and staff of related or competing organizations; and the primary users of the evaluation” (CDC, 1997).

This list may appear to be exhausting and not feasible to encompass, but the CDC reminds us that there will be variance on the levels of stakeholder involvement and how they’re involved in the process among program evaluations.

Stakeholder involvement is also an important aspect of the evaluation capacity building process (Bamberger, Rugh & Mabry, 2006; Duigan, 2003; Preskill & Russ-Eft, 2005; Sanders, 2003; Stockdill, Baizerman & Compton, 2002; Stufflebeam, 2002; Volkov & King, 2007; Wandersman et al., 2005). Stufflebeam states in his *Institutionalizing Evaluation Checklist* that an organization should “Promote and support stakeholders’ buy-in, participation, and support from all levels, e.g., by engaging representative panels to review evaluation plans and reports and working to assure that top management and governance are knowledgeable, supportive, and involved in the evaluation effort” (2002). Others focus on the importance of developing evaluation skills in stakeholders such as funders, staff, management and policy makers (Duigan, 2003; Fetterman, 2005). Volkov and King offer several examples in their *Checklist for Building Organizational Evaluation Capacity* that focus on stakeholder involvement (2007). Some of their recommended steps include to “Locate existing and enlist new evaluation champion(s) in the organization;” “Provide opportunities for sufficient input in

decision making, ensuring that people in the organization are able to use data to make decisions;” and “Establish a capable ECB oversight group (composed of members of the staff, board of directors, and community) to initiate, evaluate, and advance evaluation processes continually in the organization” (Volkov & King, 2007). Festen and Philbin include in their recommendations contacting colleagues in similar organizations to request a copy of their evaluation report to serve as an example of what can be done in your organization and as a learning piece for the evaluation committee (2007, p. 10).

Five key stakeholders likely to have an interest in the evaluation of nonprofit human service programs are included in this research – the ED of the NHSO, the Board Chair, the program staff, funders and evaluators. They have been chosen to ensure a multi-stakeholder approach for attempting to determine the factors that drive program evaluation capacity. Each of these stakeholders’ roles in the evaluation process, and more specifically in evaluation capacity building, will be discussed in the next section.

Executive Directors

The late Peter Drucker stated that “Every knowledge worker in modern organization is an ‘executive’ if, by virtue of his position of knowledge, he is responsible for a contribution that materially affects the capacity of the organization to perform and obtain results” (1967, p. 5). This definition aptly fits an executive director (ED), who assumes the highest hierarchical position, is typically the highest paid staff in a nonprofit, 501 (c) (3) organization, and some suggest is “the center of leadership for the organization” (Herman & Heimovics, 1991, p. 54). The ED plays an important role in determining their organization’s success by shaping its vision (Lynch, 1993), working with their board of directors to lead their organization’s strategic planning process (Dym

& Hutson, 2005) and ensuring the organization has the necessary financial resources to operate in an ever changing funding environment (Herman & Heimovics, 1991). The ED is ultimately responsible for the management of financial, human and capital resources the organization requires to deliver programs and work towards satisfying its mission.

EDs face the challenge of managing the political environment that has them accountable to a myriad of stakeholders (Tschirhart, 1996) including their board, staff, volunteers, consumers, community, funding organizations, accreditation organizations, government agencies and others, some of whom may have competing demands for the ED and the organization. The ED must balance these accountabilities while integrating their organization's mission, acquisition of adequate financial resources, and strategy (Herman & Heimovics, 2005). EDs must respond to both external issues such as new legislation affecting their organization and internal issues such as managing personnel (Block, 2001a). They are also responsible for expanding the boundaries of their organization through networking, representing their organization to the community, and responding to an ever changing environment (Drucker, 1990; Herman & Heimovics, 2005). An ED's efforts in boundary expansion can help fulfill the overarching expectation of overcoming organizational and environmental constraints to work towards accomplishing the organization's mission (Block, 2004).

The vast literature on leadership provides many definitions, however this study adopts the following, "...a process whereby an individual influences a group of individuals to achieve a common goal" (Northouse, 2004, p. 3). EDs of nonprofit organizations are charged with balancing being effective managers who plan, organize, staff and control operations while also effectively leading their organizations to remain

competitive and reach their full potential in an ever changing environment. In doing so an ED must combine administrative tasks such as planning and budgeting with leadership actions such as building an effectively working group to be an agent of change (Powell, 1995). There are many traits of successful leaders of nonprofit organizations including being proactive to make things happen rather than merely responding to the demands of external forces (Lynch, 1993), communicating effectively with stakeholders (Denhardt, Denhardt & Aristigueta, 2002), and aligning personality and behavior with the organization's culture (Dym & Hutson, 2005).

EDs are asked to support their organization's evaluation process Block's required competencies for EDs (Block, 2001a, pp. 103-105) are highlighted below and addressed individually for their relevance and importance in evaluation capacity building. The *planning* skills EDs use for establishing goals for their organization can also be helpful for incorporating program evaluation into the organization's strategic plan, developing an evaluation capacity building plan, planning a program or planning an evaluation. An ED's competency for *organizing* when determining necessary resources for operations, staff and programs can be helpful in placing an internal evaluator in an appropriate hierarchical position, putting together an evaluation advisory committee or linking their program's needs with the corresponding resources necessary to evaluate them. *Motivating* is a competency already covered in previous sections. Overall, if staff is not motivated to do their jobs well and feel good about their work, they will probably not be interested in participating in the evaluation process. An ED's competency for *decision making* requires adept skills in analyzing information and determining what courses to take for the organization which are analogous to the program evaluation process. The ED

can decide that program evaluation is a priority for their organization and therefore convey its importance to the board, staff, volunteers and other key stakeholders, as well as plan, budget and allocate resources for it (United Way of America, 1996). *Delegating* is a necessary competency for an ED's survival, as those leading NHSOs particularly are usually overcommitted for their time and have arguably unrealistic expectations from their board. An ED can delegate the evaluation function to dedicated staff as well as use delegation effectively to allow staff to have enough time to engage in the evaluation process.

An ED's role in *coordinating* activities among staff through horizontal and vertical integration can also include the coordination of those related to the program evaluation process. The *reporting* competency for EDs is necessary for their ability to ensure accountability among staff and the supervisors they report to. This accountability can be driven by the expected outcomes of individual employees which may be directly related to the outcomes of the organization's programs, especially if program evaluation is used as a tool for performance measurement. *Supervising* direct staff reports as well as evaluating the supervision of their subordinates includes solving operational problems, providing the necessary tools and resources for staff to do their jobs, and meet expected outcomes for performance all of which can cross over to the realm of program evaluation similar to Stufflebeam's process evaluation in his CIPP Model (2004, p. 246) (see *Funders* section for discussion). Lastly, *managing finances* and *fundraising* can ensure dedicated funding for the evaluation function, the use of evaluation results to meet external funding demands and to ensure overall fiscal health and stability for the organization which are ideal, if not necessary, conditions for program evaluation. It is

important to note that the applicability of these analogous recommendations depends largely on the organization's size, the orientation of the ED, and environmental factors.

These competencies along with their roles previously discussed in developing the organization's culture for evaluation; managing the socio-political environment of their organization and for program evaluation; encouraging, enabling, fostering and rewarding organizational learning; and ultimately deciding their organization's course for program evaluation places EDs as critical drivers for the evaluation capacity building process.

Board of Directors

The Board of Directors of an NHSO is a group of volunteers fiscally and legally responsible for the organization, as accountability ultimately resides with the board (Block, 2001b). In spite of this, there is scarce literature covering the board's involvement in or knowledge of the program evaluation process. Carman's study revealed a mix of levels of involvement and interest ranging from the more hands-on boards being interested in program evaluation to boards only caring when a funder required it or made a change in their grantee reporting process to not caring about it at all (2007). A brief review of the board's responsibilities and functions helps provide insight as to whether they relate to ECB and how the board can play a role in the ECB process.

The board hires, fires and evaluates the ED. However, rather than choosing a subordinate model of governance, some recommend a governance model where they work in partnership with the ED (Smith, Bucklin & Associates, Inc., 2000) and employ "board-centered leadership" (Herman & Heimovics, 1991) to ensure the organization is managed effectively and working towards achieving its mission. The board works toward these primary goals by fulfilling the three standards of conduct – the *duty of*

obedience, the *duty of care* and the *duty of loyalty* (Axelrod, 2005; Block, 2001b; Burgess, 1993). The duty of obedience involves acting consistently within the organization's mission, purpose and applicable laws. The board fulfills the duty of care when it acts in good faith and in the organization's best interests when making decisions by staying informed, asking critical questions and participating in governance. The duty loyalty is fulfilled when the organization's interests come first and above individual interests, and it avoids conflict of interest.

The board has several primary responsibilities while attempting to fulfill these standards of behavior. The most common primary board responsibilities cited in the literature include determining and advancing the organization's mission, setting policies, ensuring the organization has adequate resources, serving as a connection to the community through outreach, and providing oversight of management (Axelrod, 2005; Block, 2001b; Hay, 1990; Oster, 1995; Smith, Bucklin & Associates, Inc., 2000). Ensuring organizational and programmatic effectiveness emanates from the standards of behavior (duties) and these primary responsibilities in particular the determining and advancing of the organization's mission. If an organization's programs represent the organization's work towards its mission, then this translates to the evaluation of those programs. Researchers and scholars vary on how they characterize this board responsibility. They include the monitoring and assessing the organization's work towards its mission (Smith, Bucklin & Associates, Inc., 2000); the board being "...the evaluators of the accomplishment of the organizational mission and related objectives" (Hay, 1990, p. 154); and "...ensuring policies are in place to evaluate the organization's

programs and services to determine if they advance the mission and are effective...”
(Axelrod, 2005, p. 137).

Some suggest that the board should evaluate the program regularly by probing and questioning to make sure the programs are sound, consistent with the mission and with what has been promised by the organization (Burgess, 1993; Carver, 2002) while others suggest their participation be more hands-on with a role in the planning, monitoring and evaluation of programs (Scribner, 2004). O’Connell recommends that if the board or other internal stakeholders do not have the technical aptitude to evaluate their programs, then the board should hire external assistance (1985). Carver recommends that boards go beyond categorizing service delivery as a process to focus more on the difference programs make in their consumers’ lives (2002). His advice represents the switch from the outputs model which focuses more on how many consumers have been served and how many resources are necessary to deliver the services to an outcomes model which is concerned with how well those consumers are being served with respect to the intended changes in their condition, status or behavior. Carver echoes Aristotle by reminding us that ends are superior to means and that they should be considered and determined first before developing the means to support them (2002).

Program Staff

In the past, the majority of human services evaluation was conducted by academic social workers or researchers (Reamer, 1998), however we now see program evaluation has become an integral part of human service professions. Gardner states that, “Ideally, research and/or evaluation are built into practice – as they are into training – rather than

being an optional extra” (2006, p. 32). The National Organization for Human Service Education (NOHSE) and Council for Standards in Human Service Education (CSHSE) include in their set of human service worker competencies, “*Skill in planning, implementing, and evaluating interventions*” (NOHSE & CSHSE in H.S. Harris, D.C. Maloney & F.M. Rother (Eds.), 2004, p. 124). These examples of evaluation’s role in the human services are supported by the standards for the professions, four of which are selected as examples here, which contain explicit references to evaluation as part of their practice.

The National Association of Social Workers’ (NASW) Code of Ethics has a section entitled “Evaluation and Research” which contains 16 specific items pertaining to evaluation. Four have been selected as examples of the relationship between evaluation and social work practice:

- (a) Social workers should monitor and evaluate policies, the implementation of programs, and practice interventions.
- (b) Social workers should promote and facilitate evaluation and research to contribute to the development of knowledge.
- (n) Social workers should report evaluation and research findings accurately. They should not fabricate or falsify results and should take steps to correct any errors later found in published data using standard publication methods.
- (o) Social workers engaged in evaluation or research should be alert to and avoid conflicts of interest and dual relationships with participants, should inform participants when a real or potential conflict of interest arises, and should take steps to resolve the issue in a manner that makes participants' interests primary.

Code of Ethics of the National Association of Social Workers (1999)
from <http://www.socialworkers.org/pubs/code/code.asp>

The Code of Professional Ethics and Practices of the National Association of Workforce Development Professionals (NAWDP) contains the following information regarding evaluation:

1.6 Standardized Assessment Instruments - Members promote the professional and ethical use of standardized assessment instruments. Members understand that standardized instruments are a valuable part of an assessment process but may require substantial expertise to use properly. Members attempt to prevent and/or correct situations in which standardized instruments are used improperly.

1.6c. Staff Preparation - Members involved in standardized instrument selection, administration, scoring and interpretation must be competent to perform their role in the assessment process by virtue of academic preparation, in-service training, prior experience or supervised on-site training.

5.3 Program Evaluation and Research - Members contribute to the knowledge base of the workforce development profession by participating in and supporting research and other activities that identify successful strategies and programs.

NAWDP's Code of Professional Ethics and Practices (2002) from http://www.nawdp.org/AM/Template.cfm?Section=Code_of_Ethics&Template=/CM/ContentDisplay.cfm&ContentID=1358

The Code of Ethics for the American Counseling Association (ACA) has an entire section called "Evaluation, Assessment, and Interpretation." Its introduction reads, "Counselors use assessment instruments as one component of the counseling process, taking into account the client personal and cultural context. Counselors promote the well-being of individual clients or groups of clients by developing and using appropriate educational, psychological, and career assessment instruments" (ACA, 2005, p. 11).

Some examples in this section for the counseling practice's role in evaluation include:

E.2.c. Decisions Based on Results - Counselors responsible for decisions involving individuals or policies that are based on assessment results have a thorough understanding of educational, psychological, and career measurement, including validation criteria, assessment research, and guidelines for assessment development and use.

E.12. Assessment Construction Counselors use established scientific procedures, relevant standards, and current professional knowledge for assessment design in the development, publication, and utilization of educational and psychological assessment techniques.

ACA's Code of Ethics (2005) from
<http://www.counseling.org/Resources/CodeOfEthics/TP/Home/CT2.aspx>

The American Psychological Association's (APA) Ethical Principles of Psychologists and Code Of Conduct has an entire section entitled "*Assessment*" that contains the following examples of the psychology profession's relationship with evaluation:

- (a) Psychologists administer, adapt, score, interpret, or use assessment techniques, interviews, tests, or instruments in a manner and for purposes that are appropriate in light of the research on or evidence of the usefulness and proper application of the techniques.
- (b) Psychologists use assessment instruments whose validity and reliability have been established for use with members of the population tested. When such validity or reliability has not been established, psychologists describe the strengths and limitations of test results and interpretation.
- (c) Psychologists use assessment methods that are appropriate to an individual's language preference and competence, unless the use of an alternative language is relevant to the assessment issues.

APA's Ethical Principles of Psychologists and Code Of Conduct (2002)
from <http://www.apa.org/ethics/code2002.pdf>

The American Camping Association has 14 essential areas of camp management, two that are related to evaluation. One is "Program Design & Activities" that covers "Designing and evaluating the effectiveness of an organizational structure for the camp program that is appropriate for persons being served, the camp's philosophy, goals and objectives, and the environment utilized" (American Camping Association, 2008). Some of the programs it covers include Nature & Environmental Awareness; Outdoor Living Skills; Skits, Stunts, & Special Events; Sports; and Stories, Songs, & Memories.

The other essential area is “Mission & Outcomes” that covers “Identifying and articulating the mission and philosophy to parents, participants, board members and funders, and evaluating current issues and their implications on desired goals and outcome objectives” (American Camping Association, 2008). It has a section called “Achieving Quality Outcomes” that helps camps answer the questions about what constitutes a quality program, how we would recognize one, and what measures help determine quality. Their ultimate question for the outcomes of camping programs is “How can we be assured that the programs we develop are designed to appropriately meet developmental needs that provide youth with resilience and help them become productive as adults?” (American Camping Association, 2008).

These specific professions only represent a portion of all found in human services. There is also the generalist practice which, as its name implies, incorporates a broad knowledge base, range of skills and professional values for addressing various human social problems found in different systems of service delivery (Hull, Jr. & Kirst-Ashman, 2004). Generalist practice differentiates itself with the aforementioned specialized professions in that it address multiple issues based on a philosophy that human social problems are complex and are better dealt with holistically or within a continuum of service delivery. Human service generalist practitioners play multiple professional roles that are situation and context dependent. The other primary processes for this practice that Hull, Jr. and Kirst-Ashman are similar to those of other human service professions and they include “working effectively within an *organizational structure* and doing so under supervision;” applying “*critical thinking skills to the planned change process*,” and emphasizing *client empowerment*” (Hull, Jr. & Kirst-Ashman, 2004, p. 4).

NHSOs that incorporate generalist practice include those where a practitioner delivers the multiple services that have been determined necessary to address a particular human social problem. For example, in a job readiness program, he or she would help the consumer with skills for writing a resume, job searching, job interviewing and possibly assist with job placement. This helps describe the Generalist Intervention Model (GIM) which incorporates six steps to the change process that include engagement, *assessment, planning, implementation, evaluation and termination* (Hull, Jr. & Kirst-Ashman, 2004). The sequence and flow of these steps in GIM visually depict its logic model and describe what we might consider program theory.

The first stages of engagement and assessment establish the relationship between the consumer and practitioner and help determine the consumer's specific problems that require attention. Planning represents the planning of the intervention by turning the consumer's problems into needs; develop goals for satisfying those needs; and evaluating the various types, levels, pros and cons of the interventions intended to attain the goals. The emphasis for the clarification of goals in the generalist practice literature is analogous to the importance of clarifying goals and objectives for programs in program evaluation. Implementation is simply the delivery of the intervention. Evaluation is the assessment of the impact of the interventions in terms of whether they attained their goals. GIM in being consistent with its valuing the empowerment of consumers recommends that they be involved in the evaluation process as much as possible. Termination represents the inevitable part of the service delivery continuum as all practitioner-consumer relationships end at some point. The evaluation of the

interventions helps practitioners and consumers make that decision jointly (Hull, Jr. & Kirst-Ashman, 2004).

In summary we can see from the information from these human service professions that evaluation is not portrayed as a function necessarily to be performed by someone other than the practitioner and that the function is a core component of their practice. We can also see in some of the language how evaluation is described as the process to assess the effectiveness of the intervention on the consumer as well as how evaluation contributes to each profession's body of knowledge. One example of such a contribution a practice's body of knowledge has been program staff's growing interest in contributing to human services literature and presenting their work at professional conferences (Reamer, 1998).

Program staff's role in the program evaluation process is a major factor in building evaluation capacity at the individual level through professional standards, education, training and the experience of their practice. This role also contributes to evaluation capacity building at the organizational level when they establish, along with the evaluator if they are not conducting the evaluation alone, answers to the four important questions for beginning the evaluation process. A determination should be made for 1) who has requested the evaluation and why; 2) what the goals are for the program evaluation; 3) how the results will be used; and 4) who will be designing and carrying out the evaluation (Reamer, 1998, pp. 273-276).

Program staff are arguably the most critical stakeholders to work with evaluators in the effort to evaluate programs. Their intimate knowledge of the program, experience, and knowledge for their profession all are important aspects of the program evaluation

process. Program staff may be solicited to serve on the evaluation advisory committee because of these characteristics. Their buy-in and support for program evaluation impacts their relationships with evaluators. If this is a productive, cooperative, and collaborative working relationship, the chances of producing a high quality evaluation are increased. More specifically, if program staff serve as “evaluation champion(s) in the organization,” engage in evaluation activities, and are “supportive of change,” they can significantly contribute to the evaluation process (Volkov & King, 2007). Conversely, if the relation is unenthusiastic and uncooperative, the chances are likely the evaluation will not be high quality or even possible.

Funders

Funders such as government agencies, community foundations, the United Way, independent foundations, corporate foundations, and corporate giving programs are also consumers of program evaluation in that they either call for it and/or use the results. When they fund nonprofit human service programs they likely will have some interest in how their funds were used for the program, how efficient and effective those programs are at addressing a particular social problem, and possibly will use such information for funding decisions (Monette, Sullivan & DeJong, 2002; Tripodi, 1983). These entities have specific interests and priorities for the funding and/or the evaluation of programs that sometimes move in different directions. This presents a challenge for NHSOs especially those who rely on funding from multiple sources, and for evaluators (Bernstein, Whitsett & Mohan, 2002).

Rooney and Frederick’s study that analyzed foundations’ overhead funding policies and their impact on nonprofit organizations included 389 educational

organizations and 491 NHSOs (2007, p. 29). Program evaluation was not specifically one of the seventeen costs which included rent, office supplies, strategic planning, accounting fees, and fundraising among others. However, it is important to note that of the entire sample of 710 foundations, 69% of all foundations, 71.9% of the independent foundations, 70.8 of the corporate foundations, and 61.7 of the community foundations reported funding all seventeen costs (2007, p. 21). The amount of annual giving and the age of the foundations had a relationship with the funding of these expenses. Foundation giving more than \$6.5 million a year were more likely to fund all types of these costs. Foundations between ten and fifty-five years of age were less likely to fund all costs compared with the younger foundations of ten years or less of age (2007, p. 12). Foundations funding NHSOs were more likely to fund all costs than those foundations funding public-society benefit organizations (2007, p. 12).

Nonprofit organizations have reported their financial information based on the assumption that their ability to acquire funding will depend on spending as little as possible on administrative or overhead expenses. One former leader of a large funder in the Atlanta MSA incorporated his “ten percent rule” which restricted his organization, and strongly advised grantee organizations, to having ten percent or less than the organization’s overall expenses be administrative, leaving ninety percent or more to be program related expenses. The Rooney and Frederick study helps to dispel this assumption by indicating that “foundations do fund nonprofits’ overhead expenses, but they do so mostly through the inclusion of overhead expenses within program grants” (2007, p. 17). This raises important questions for the funding of evaluations of nonprofit human service programs. First, is program evaluation typically categorized as a program

cost or administrative cost? Second, is program evaluation typically included in program grants regardless of cost category? Third, if program evaluation is part of the program grant, are such funds that foundations provide for this cost adequate to pay for the evaluation? These answers to these questions are an important part of the resources aspect of ECB.

Another challenge funders impart to their NHSO grantees relates to how they influence the departmentalization of services based on their funding priorities and funding streams. This particularly impacts human service professionals who attempt to employ a flexible, holistic approach to addressing human social problems (Gardner, 2006). On an operational level it affects how consumers go through the in-take process, how their progress is assessed, and how they may be referred to other NHSOs providing the other services they need. This can place difficult challenges for consumers who have to see multiple human service professionals, go through multiple administrative processes all while dealing with scheduling, family and transportation issues. The ultimate consequences for the departmentalization of services can be that some consumers may not be served well, through the necessary duration to address their problems, or at all.

The implication for human service workers is that it prevents them from treating the consumer as a whole person which is a core principle for their profession (Gardner, 2006). Funders, especially government agencies, might consider creative ways to support a more holistic approach to solving complex social problems. The majority of funding for NHSOs is oriented for treatment programs instead of prevention programs (Gardner, 2006), which suggests that funders might want to focus more on the causes of social problems and fund their prevention which will likely be less costly than their treatment,

or consider balancing prevention with treatment. The human services sub sector has examples of organizations geared towards preventing child abuse, drug addiction, pregnancy, domestic violence and other social issues. There are more examples today than in previous decades of private and public funders encouraging and funding collaborative efforts among NHSOs to more broadly address social problems.

Some scholars and researchers have suggested strategies NHSOs can employ to contend with these challenges. They include acquiring a better understanding of the various interests of and relationships between stakeholders through network analysis, building an evaluation culture within the organization so that stakeholders are committed to evaluation (Fredericks, Carman & Birkland, 2002); incorporating various evaluation methods to better fit the needs of stakeholders (Tang, Cowling, Koumjian, Roesler, Lloyd & Rogers, 2002); communicating effectively with stakeholders by explaining what evaluation is, how it can contribute to service delivery, outcomes and management and that it is not for grading or punishment (Fredericks, Carman & Birkland, 2002); negotiating with stakeholders including the sponsoring funder (Guzman & Feria, 2002) which presents an opportunity for a NHSO to engage in dialogue with their funder and help balance the imbalance of power in their relationship.

Others suggest more explicit funder support for evaluation beyond relationships and dialogue. A funder's role in ECB can be to ensure that capacity building is expected of the evaluation, support ECB efforts by providing additional experts for technical assistance as they are needed, and share their expertise in managing programs and raising funds (Fetterman, 2005). Funders are identified as one of the three primary stakeholder groups driving empowerment evaluation along with evaluators and the evaluation

stakeholder community. Their roles in empowerment evaluation's principles represent more involvement than what would typically be expected of funders of nonprofit, human service programs. Some of them for select principles serving as examples of this are depicted below in Table 3:

Table 3 – Funder Roles in Empowerment Evaluation

EMPOWERMENT EVALUATION PRINCIPLE	FUNDER ROLE
Improvement	<ul style="list-style-type: none"> • Provides the financial support required of a stakeholder community engaged in improvement-oriented evaluation efforts • Rolls up sleeves and helps problem-solve to improve the program
Community Ownership	<ul style="list-style-type: none"> • Respects the autonomy of the organization or agency to pursue the evaluation, as deemed appropriate (in conjunction with the evaluator) • Encourages institutional ownership of the evaluation • Supports institutionalization of evaluation in the organization
Inclusion	<ul style="list-style-type: none"> • Encourages the stakeholder community and the evaluators to be as inclusive as possible
Democratic Participation	<ul style="list-style-type: none"> • Supports democratic participation with appropriate funding and an appreciation for the additional time required
Evidence-Based Strategies	<ul style="list-style-type: none"> • Shares evidence-based strategies that have been successful in similar funded programs

Extracted and adapted from Table 3.1, pages 55-72 in Fetterman, D. M. (2005). *Empowerment Evaluation Principles in Practice*. In D.M. Fetterman & A. Wandersman (Eds.), *Empowerment Evaluation Principles in Practice*. New York: The Guilford Press.

Hawkins' study of community, corporate, and independent foundations supported this notion and yielded recommendations for them to employ. One is for foundations to employ evaluators on their staff that can help build evaluation capacity within their organization as well as their grantees, and free up time for program officers to focus on other issues. Another is for foundations to come closer to acquiring more valid and reliable information on the programs they fund by developing a comprehensive evaluation strategy based on Stufflebeam's Context, Input, Process & Product (CIPP)

model (Hawkins, 1984). “The CIPP model is a comprehensive framework for guiding formative and summative evaluations of projects, programs, personnel, products institutions, and systems” (Stufflebeam, 2004, p. 245). The model has been adapted into Table 4 shown below from Stufflebeam’s text for the purpose of summarization.

Table 4 – CIPP Model for Evaluation

EVALUATION	ASSESSSES	MICRO FOCUS (For)	MACRO PURPOSE (To)
Context	Needs, problems, assets, and opportunities	Helping decision makers define goals and priorities	Help the broader group of users judge goals, priorities, and outcomes
Input	Alternative approaches, competing action plans, staffing plans and budgets	Feasibility and potential cost-effectiveness	Met targeted needs and achieve goals
Process	The implementation of plans	Helping staff carry out activities	Help the broad group of users judge program performance and interpret outcomes
Product	Outcomes – intended and unintended, short-term and long-term	Helping staff keep an enterprise focused on achieving important outcomes	Help the broader group of users gauge the effort’s success in meeting targeted needs

Adapted from: Stufflebeam, D. (2004). The 21st-Century CIPP Model. In M.C. Alkin (Ed.), *Evaluation Roots* p. 246. Thousand Oaks, CA: Sage Publications, Inc.

Foundation use of this model for engaging in supporting evaluations of their grantees’ programs should benefit in building strong relationships with their grantees (Patton, Bare & Bonnet, 2004). This can lead to such relationships resembling more like partnerships where funders and grantees have mutual goals. Funders can also have a better understanding of their grantees’ program needs, evaluation needs, the intended outcomes of the program, and how the effectiveness of the grantee’s program impacts the effectiveness of their grant making process and goals for social betterment.

Private foundations have been demanding more comprehensive program information from their grantees (Festen & Philbin, 2007). Also, more are moving towards asking for program evaluation up front, so their grantees are not surprised at the end of the program cycle and wind up scrambling to conduct an evaluation they are not prepared for. Some of the more prominent foundations that have supported program evaluation over the past decade include Annie E. Casey Foundation, Bill and Melinda Gates Foundation, Robert Wood Johnson Foundation, W. K. Kellogg Foundation, and the David and Lucille Packard Foundation. These foundations usually incorporate evaluation as part of their funding process and relationships with their grantees, have information about evaluation available on their web sites, and have participated in the American Evaluation Association's Nonprofit and Foundation Evaluation Topical Interest Group (TIG) (American Evaluation Association, 2001). The TIG was formed in 2001 and membership has almost doubled since then to 891 members (Nonprofit and Foundation Evaluation TIG, 2006).

Evaluators

Evaluation has experienced some of the milestones used as criteria for fields becoming professions while progress through the stages of their lifecycle such as becoming a full-time occupation, creating a training school, having courses available at a university, founding the first local professional association, founding a national professional association, and developing a code of ethics (Rothman, 1987). The evaluation profession in the United States has especially evolved over the last 20 years when it experienced in 1986 the founding of the American Evaluation Association (AEA), "an international professional association of evaluators devoted to the application

and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation” that currently has approximately 5600 members across the U.S. and 60 other countries (AEA, 2008). The association has seen a 23% increase in membership from 2006-2008 and a 13% increase in annual conference attendance between 2006 and 2007. AEA currently has 24 local evaluation associations (affiliates) in 21 states and Puerto Rico (AEA, 2008). AEA has topical interest groups (TIGs) focused around a special topic or interest to subgroups of the association. The two most relevant to this research include the Nonprofit and Foundation TIG founded in 2001 with currently over 900 members and the Organizational Learning & Evaluation Capacity Building TIG founded in 2006 with currently over 700 members.

The evaluation profession now has two U.S.-based journals, *New Directions for Evaluation* and the *American Journal of Evaluation*, and AEA reports that there are approximately 38 universities in the U.S. that “...offer graduate programs or certificate programs either directly in evaluation or with available concentrations in evaluation” (AEA, 2008). In addition to these academic instruction opportunities, evaluation academics and practitioners can gain professional development through attendance at AEA’s annual conference, The Evaluators’ Institute (TEI) which offers four different certificate programs for professional evaluators (The Evaluators Institute, 2008), or the annual AEA/CDC Summer Evaluation Institute. Other important indications of evaluation’s progress as a profession include the establishment of the program evaluation standards in 1994 (Joint Committee on Standards for Educational Evaluation, 1994) guiding principles for evaluators in 2004.

In spite of the evaluation profession's evolution and gain in prominence, as well as the increased awareness of evaluation among nonprofit professionals, there still exist images of who evaluators are and what they do that may lean towards stereotypes or misperceptions. For example, some view evaluators as merely statisticians who deal only with numbers. Some feel they represent "big brother" watching them ready to punish them by providing information that will lead to the elimination of their program. Others believe they conduct evaluations of programs by themselves without any need of assistance from program staff or other stakeholders. Still others view evaluators as people who are hired guns to provide a tangible deliverable primarily to satisfy an external demand. These images take away from the value evaluators can contribute through their various roles in the program evaluation and evaluation capacity building processes.

These images may also not be accurate due to the changing environment in which evaluators operate. The roles of evaluators have expanded and evolved over time to where they can no longer just be the experts in evaluation tools, methods and statistics. This is driven by the fact that they "...work in highly complex, cultural, social and political environments and interact with a wide range of stakeholders" (Preskill & Catsambas, 2006, p. 42). Guba and Lincoln in 1989 provided four examples of how evaluators' roles had evolved in a broader sense that are still relevant today. First, the evaluator shifted from being more of a controller to a collaborator based on the increasing recognition and use of participatory evaluation and stakeholder involvement methods and techniques. Second, the evaluator extended beyond being an investigator to become a teacher and a learner. This resulted from increasing recognition of social constructions

among stakeholders, fostering of organizational learning and engaging in reflective practice. Third, the evaluator evolved from being a discoverer to become a shaper of reality. Evaluators now have to go beyond just reporting results to actively participate in helping stakeholders shape the reconstruction of reality for the program. Lastly, they suggested “the evaluator divests him- or herself of the role of passive observer and recognizes and embraces the role of change agent (Guba & Lincoln, 1989, pp. 260-261).

The advent of program evaluation standards, guiding principles for evaluators and an emphasis on ethical behavior have also contributed to this evolution. All of these influences together account for what Preskill and Catsambas indicate are today’s expectations for evaluators:

- Design and implement the evaluation in culturally responsive and appropriate ways
- Be more transparent
- Be performance-improvement and learning oriented
- Practice evaluation in the context of continuous change, that is, where the program being evaluated continues to evolve (2006, p. 41)

This section discusses the various specific roles evaluators play and activities they engage in for the evaluation capacity building and program evaluation processes.

Evaluators must have specific core competencies if they are to successfully assume these roles and engage in these activities. Sanders indicated that they include “...the ability to describe the object and context of an evaluation; to conceptualize appropriate purposes and frameworks for the evaluation; to identify and select appropriate evaluation questions, information needs, and sources of information; to select means for collecting and analyzing information; to determine the value of the object of an evaluation; to communicate plans and results effectively to audiences; to manage the evaluation; to

maintain ethical standards, to adjust for external factors influencing the evaluation; and to evaluate the evaluation (metaevaluation)” (1979, as cited in Fitzpatrick, Sanders & Worthen, 2004, p. 7). Stevahn, King, Ghore and Minnema in 2001 created their taxonomy for essential evaluator competencies (revised in 2005), and the main categories include Professional Practice, Systematic Inquiry, Situational Analysis, Project Management, Reflective Practice and Interpersonal Competence⁸ (2005, pp. 45-51). Their justification for this effort includes the fact that there is no licensing or credentialing for evaluators; consumers of evaluation do not have a guide to help them make decisions for hiring evaluators based on qualifications; similarly those interested in becoming evaluators do not have a guide which lets them know what they need to study and what skills they need to acquire; university and professional development programs can now have more consistency for developing curriculum; and lastly the concern that the profession mostly has developed models for practice based on specific contexts in lieu of theory-based models that tested and validated for effective practice (Stevahn, King, Ghore & Minnema, 2005).

Many roles evaluators play, some multiple and simultaneously, include educator, facilitator, coach, mentor, technical advisor, trainer (Chen, 2005; Patton, 1997; Preskill, 2004; Preskill & Catsambas, 2006), methodological expert, judge (Guba & Lincoln, 1989; Patton, 1997) and change agent (Brun, 2005; Guba & Lincoln, 1989; Patton, 1997; Sonnichsen, 1994). One important role an evaluator can play is to work towards ensuring utilization of the evaluation results (Bell, 1994; Patton, 1997). This role involves the *utility* standards for evaluation which “...are intended to ensure that an evaluation will serve the information needs of intended users” (Joint Committee on Standards for

Educational Evaluation, 1994). These standards include strategies, responsibilities and tasks such as collecting information related to the interests of consumers, writing clear reports so stakeholders can understand the information, and disseminating findings to stakeholders in a timely fashion. The Guiding Principles for Evaluators also contain information concerning the evaluator's role in encouraging the use of results such as to "Negotiate honestly with clients and relevant stakeholders concerning the costs, tasks, limitations of methodology, scope of results, and uses of data" and "Be explicit about their own, their client's, and other stakeholder's interests and values related to the evaluation" both of which reside in the principles for *Integrity/Honesty* (American Evaluation Association, 2004).

Ensuring stakeholder involvement is an important role for evaluators, and *engaging stakeholders* is the first step in the Centers for Disease Control and Prevention's (CDC) Framework for Program Evaluation in Public Health (CDC, 1997). This can involve assessing together with the representatives of the NHSO who are the stakeholders who are "involved in program operations...served or affected by the program...and primary users of the evaluation" (CDC, 1997). Some scholars and practitioners suggest being as expansive and inclusive as possible (Abma, 2006; Bamberger, Rugh & Mabry, 2006). If this step does not take place in such a holistic fashion, there is chance that important stakeholders will be left out of the process and their voices will not be heard. There is also a chance that stakeholders' needs will not be completely understood, so evaluators can play an important role in ensuring democratic participation (Fitzpatrick, Sanders & Worthen, 2004). Stakeholder involvement in this

regard can be an effective means for managing the socio-political environment for program evaluation by reducing uncertainty and confusion (Chen, 2005).

Fitzpatrick, Sanders and Worthen offer some issues evaluators are recommended to seek answers for that will help them navigate the socio-political context for program evaluation. They address assessing levels of cooperation from stakeholders and whose cooperation is required; who has power within the context; how should the evaluator relate to the stakeholders i.e. as impartial outsider, consultant, facilitator, advocate or some combination of these; who is invested in the evaluation and who stands to gain or lose depending in the scenario; who the contact is that needs to be informed and updated; and what risk management and legal issues require policies or procedures (Fitzpatrick, Sanders and Worthen, 2004, p. 216). They also suggest that evaluators follow the two most important principles for dealing with political pressures by “avoiding conflict of interest and insisting on open, fair, and complete disclosure of findings” (2004, p. 437).

Stimulating dialogue will help maximize stakeholder involvement and work towards a democratic decision making process (Fitzpatrick, Sanders & Worthen, 2004). One of the guiding principles for *Responsibilities for General and Public Welfare* states that an evaluator should “Include relevant perspectives and interests of the full range of stakeholders” (American Evaluation Association, 2004). One of the biggest challenges evaluators face is the stimulating and nurturing of stakeholder support and buy-in for the evaluation process. Fostering stakeholder involvement can contribute to a more comprehensive understanding of evaluation and increase their commitment to the process and using results (Preskill & Catsambas, 2006). One recommended strategy is forming an evaluation work group or advisory group comprised of a cross-section of stakeholders

potentially including program staff, the ED, a board member, a program consumer, a direct service volunteer and/or a policy group representative (Brun, 2005; Mark, Henry & Julnes, 2000; Pawlak & Vinter, 2004; United Way of America, 1996; Volkov & King, 2007; Wholey, 1994).

An important role that evaluators play in the program evaluation and evaluation capacity building processes can be easily overlooked if they are called into the process after the program is delivered. Evaluators can assist stakeholders in planning programs by assisting them in developing the program's theory and/or logic model (Chen, 2005; Patton, 1997). They can also help them work through the rationale for the program and corresponding evaluation. This process has been called *developmental evaluation* which Patton defines as "evaluation processes undertaken for the purpose of supporting program, project, staff and/or organizational *development*, including asking evaluative questions and applying evaluation logic for developmental purposes" (1997, p. 105). The evaluator's role in developmental evaluation can comprise a collaborative effort among stakeholders to conceptualize, design and test programs.

These roles are driven by the necessary interaction with various stakeholders and are based on special interpersonal skills. They include, but are not limited to, negotiation (Fitzpatrick, Sanders & Worthen, 2004; Guba & Lincoln, 1989; Guzman & Feria, 2002; Patton, 1997; Preskill & Catsambas, 2006), asking probing, thought-provoking and critical questions, being a good listener, being culturally competent and sensitive to cultural differences among stakeholders (Preskill & Catsambas, 2006), conflict resolution, and facilitating dialogue and understanding organizational change and learning (Patton, 1997; Preskill & Catsambas, 2006; Russ-Eft & Preskill, 2001).

Evaluators can engage in various activities and perform various tasks depending on the context, their specific role in the agreement or contract for evaluation, and whether they are an internal evaluator as staff for the organization or an external evaluator as a consultant. Some examples of these activities may include developing the rationale for conducting the evaluation, collecting, analyzing and reporting information, forming evaluation advisory teams, communicating regularly with stakeholders, making presentations to the board, staff and external stakeholders, recruiting and hiring other personnel to work on the evaluation, managing budgets, developing and/or negotiating contracts (Fitzpatrick, Sanders & Worthen, 2004).

Evaluators also conduct evaluability assessments which have been previously discussed. Wholey offers six key steps for evaluators conducting evaluability assessments:

1. Involve intended users of evaluation information.
2. Clarify the intended program from the perspectives of policymakers, managers, staff, and other key stakeholders.
3. Explore program reality, including the plausibility and measurability of program goals and objectives.
4. Reach agreement on any needed changes in program activities or objectives.
5. Explore alternative evaluation designs.
6. Agree on evaluation priorities and intended uses of information on program performance.

(1994, p. 18)

These steps help determine the social context for program evaluation which comprises the social environment conditions necessary for effective evaluation and likelihood of use of results (Warren, 1974). We see in these steps repeatedly the importance of stakeholder involvement, reminding us that evaluators cannot do their jobs sufficiently by themselves. We also see the role evaluability assessments have in adhering to the

feasibility and utility standards for program evaluation within the program planning process. Evaluators that conduct evaluability assessments ultimately help stakeholders improve their program (Chen, 2005).

Financial resources can be another criterion for evaluability assessments, as they are necessary to conduct evaluations of programs. Volkov and King recommend that organizations “Assure long-term fiscal support from the board or administration—explicit, dedicated funding for program evaluation activities” (2007). Evaluators when conducting evaluability assessments at organizational and program levels can determine the limitations and flexibility of the budget for program evaluation, and at times are the stakeholder responsible for proposing the budget (Fitzpatrick, Sanders & Worthen, 2004).

Some of the costs involved include, but are not limited to, compensation for the evaluator and staff contributing their time to the evaluation process, communications, data processing requirements such as hardware or software, supplies, and equipment (Fitzpatrick, Sanders & Worthen, 2004). This helps prevent the NHSO from overlooking staff time as an expense and helps them decide how they will pay for these expenses in terms of their general operating budget versus a dedicated budget or line item for evaluation. We might hypothesize here that if an organization explicitly budgets for evaluation in their financial statements for all stakeholders to see, it indicates evaluation has been acculturated into their organization as a basic assumption rather than something conducted ad hoc only for when external stakeholders require it or resources are available.

An evaluator can use strategies that can help the organization reduce costs for, while not comprising the quality of, the evaluation. Some of these strategies will likely

be familiar to NHSOs as they attempt to remain fiscally healthy and be good stewards of resources. An evaluator can help the NHSO use volunteers or low-cost workers, local specialists to reduce travel costs, borrow equipment, acquire in-kind donations, use existing data, use inexpensive software for data collection such as an Access database, or latch on to other similar studies being conducted (Fitzpatrick, Sanders & Worthen, 2004). This part of an evaluator's role in ECB assists in maintaining the feasibility standards for program evaluation which "are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal" (Joint Committee on Standards for Educational Evaluation, 1994).

One of the most critical aspects of an evaluator's role in the evaluation capacity building and program evaluation processes is their relationship with program staff. Program staff, including program managers, are the primary contacts for an evaluator because they engage in making decisions about the program as well as operational procedures and the allocation of resources (Bamberger, Rugh & Mabry, 2006). The benefits of an effective and cooperative working relationship between these two stakeholders include increasing the chance evaluation results will be used (Patton, 1997; Sonnichsen, 1994), organizational learning will occur through the staff's acquisition of evaluative knowledge and skills (Patton, 1997; Preskill & Torres, 1999; Russ-Eft & Presskill, 2001), and overall ownership and buy-in for the evaluation process will be enhanced (Sonnichsen, 1994). Other benefits include helping to assure program staff cooperates and provides accurate information necessary to carry out the evaluation (Boulmetis & Dutwin, 2005), exposing potential or existing conflicts of interest between

stakeholders (Brun, 2005), and the joint development of the program's rationale and theory (Thomas, 2005).

These benefits will only be realized if this relationship is established at the beginning of the evaluation process (Sonichsen, 1994). Evaluators should consult program staff for information, insight, access, and perspectives. This does not compromise the evaluator's role as the expert in evaluation but rather reinforces the notion that program staff are the experts for the program. These are the people on the front lines delivering the programs and working directly with consumers. They are able to see firsthand how the program operates and construct their perceptions of the program through their lens and based on their experiences. Program staff may also have a role in conceptualizing, planning and developing the program enabling them to provide further insight for the evaluation process.

Potential challenges for this relationship to be mindful of include the staff's level of enthusiasm for the program and/or the evaluation process, their education level, knowledge about evaluation and cooperation with other stakeholders invested in the program and the evaluation process (Patton, 1997). These are challenges that should be fleshed out in the beginning of the evaluation process, possibly as part of an evaluability assessment. It is important to note that these potential challenges and others are not limited to external evaluators. Internal evaluators can suffer from not being a prophet in their own land. They may be forced to instill enthusiasm among co-workers and convene them to be active stakeholders in the evaluation process. Other challenges relevant to internal evaluators include requests from senior management for information to be used for public relations, becoming the go-to person or "dumping ground" for any work

involving assessment, being excluded from the major decisions made in the organization, and moving their organization toward using the evaluation results (Patton, 1997).

Internal evaluators may find themselves being salespersons, educators, advocates and psychologists at any time during the evaluation process. If senior management doesn't understand, value, and/or support evaluation they are unlikely to support learning opportunities for their internal evaluator such as workshops, conferences, or academic instruction thereby stunting his or her professional development. These challenges may be magnified if their organization has not experienced much evaluation (Russ-Eft & Preskill, 2001).

A review of the literature discussing recommendations for a successful evaluator-program staff relationship reveals common themes such as trust, respect, inclusion, participation, open and honest communication, empathy and ownership (Abma, 2006; Bamberger, Rugh & Mabry, 2006; Brun, 2005; Fetterman, 2005; Malloy & Yee, 2006; Patton, 1997; Preskill & Torres, 1999; Sonnichsen, 1994). These characteristics for this relationship also help contribute to less staff turnover and overall organizational instability. Overall, a successful evaluator-program staff relationship can yield long-term effects in a NHSO that contribute to the resources, structures, contexts and cultural aspects of evaluation capacity building.

Methodology

Research Questions

This study seeks to understand what factors determine the capacity to evaluate a nonprofit human service program, how they relate to and impact each other, and how they are impacted by the five participating stakeholder groups. These primary research questions are supported by several secondary research questions. What motivates stakeholders to engage in and/or support program evaluation? How do these motivations impact capacity? What specific actions have stakeholders taken to be successful in building capacity for program evaluation?

The answers to these questions will be sought after by meeting the objectives of this study which are 1) to describe what is driving an organization's efforts, or lack of, to build capacity for program evaluation; 2) to describe what organizations have done to successfully build capacity for program evaluation; 3) to summarize the perspectives for and actions taken related to program evaluation from five stakeholder groups – executive directors, board chairs, program staff, funders and evaluators; 4) to provide helpful information these stakeholder groups can use to build program evaluation capacity; and 5) to propose recommendations and new studies based on the findings. Qualitative interviews with representatives from each of the five stakeholder groups accompanied by some NHSO document review, subject to availability, will serve as the means to accomplish these research objectives.

Sampling

The complexity and interdisciplinary nature of this topic suggests a multi-stakeholder research approach be utilized to provide a variety of perspectives from the

stakeholders possibly invested in program evaluation. These sample populations include – EDs, Board Chairs, program staff, funding organizations (referred to as funders), and evaluators of NHSOs within the 20-county metropolitan statistical area (MSA) of Atlanta, Georgia and the 9-county Indianapolis, Indiana MSA. This exploratory, qualitative study includes 126 one-on-one, face-to-face interviews with these stakeholders as shown in Table 5 below:

Table 5 – Sampling of Stakeholder Groups

STAKEHOLDER GROUPS	SUB-GROUPS	# INTERVIEWS		TOTALS
		<u>GA</u>	<u>IN</u>	
Executive Directors (42)		21	21	42
Board Chairs (20)		10	10	20
Funders (24)	United Ways	2	2	4
	Community foundations	2	2	4
	Corporate foundations	2	2	4
	Independent foundations	2	2	4
	Corporate Giving Programs	2	2	4
	Government agencies	2	2	4
Evaluators (20)		10	10	20
Program staff (20)		10	10	20
TOTALS		63	63	126

The methodology for this research that follows is documented in detail, so that it may be replicated according to the guidelines for research from the American Association for Public Opinion Research (AAPOR)⁹

Sampling Procedures

Twenty-one organizations each from the 20-county metropolitan statistical area (MSA) of Atlanta, Georgia and the 9-county Indianapolis, Indiana MSA were randomly selected for a total of forty-two NHSOs. GuideStar, an online database containing information on public charities including their Internal Revenue Service (IRS) 990 forms, was used to develop the database of NHSOs for the two populations from which the samples were drawn.¹⁰ The sample of forty-two NHSOs was stratified by GuideStar's top five organizational income levels (\$250,000-\$499,999; \$500,000-\$999,999; \$1million-\$4,999,999; \$5million-\$19,999,999; \$20million)¹¹ which are taken from line 12 of their IRS 990 form, and the percentage of geographic spread within their MSA population by county. In addition to being NHSOs in the appropriate counties and income levels, organizations selected for the sample had to have filed a Form 990 within the past two fiscal years to ensure consistency and account for timing and overlap of when the research is conducted with the varying fiscal years of the organizations.¹²

The Atlanta NHSO sample is broken out by county based on how they are represented in the population within the 20 county MSA in Figure 6 while Figure 7 depicts the same comparison for the Indianapolis MSA. Groves reminds us that "A sampling frame is perfect when there is a one-to-one mapping of frame elements to target population elements. In practice, perfect frames do not exist; there are always problems that disrupt the desired one-to-one mapping" (Groves et al., 70).

Figure 6 – Comparison of Sample to Population by County (Atlanta MSA)

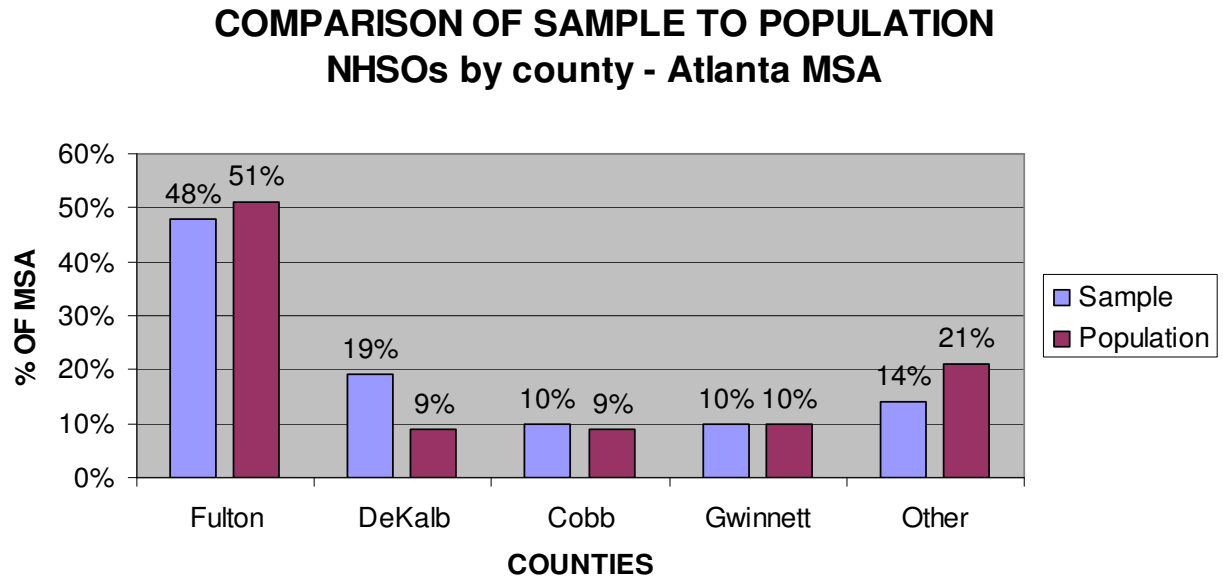
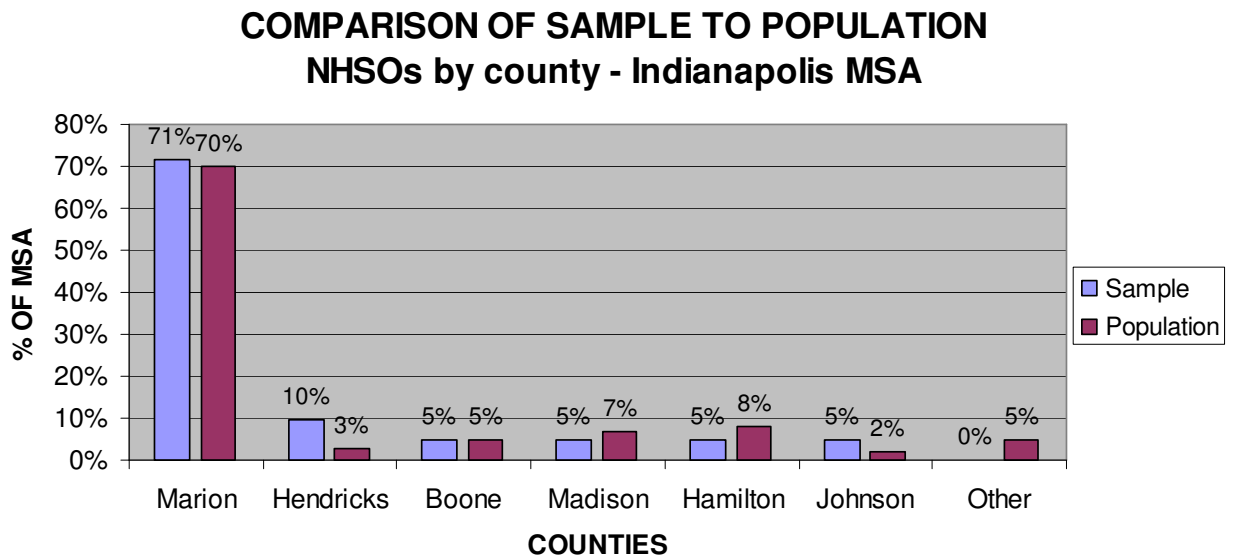


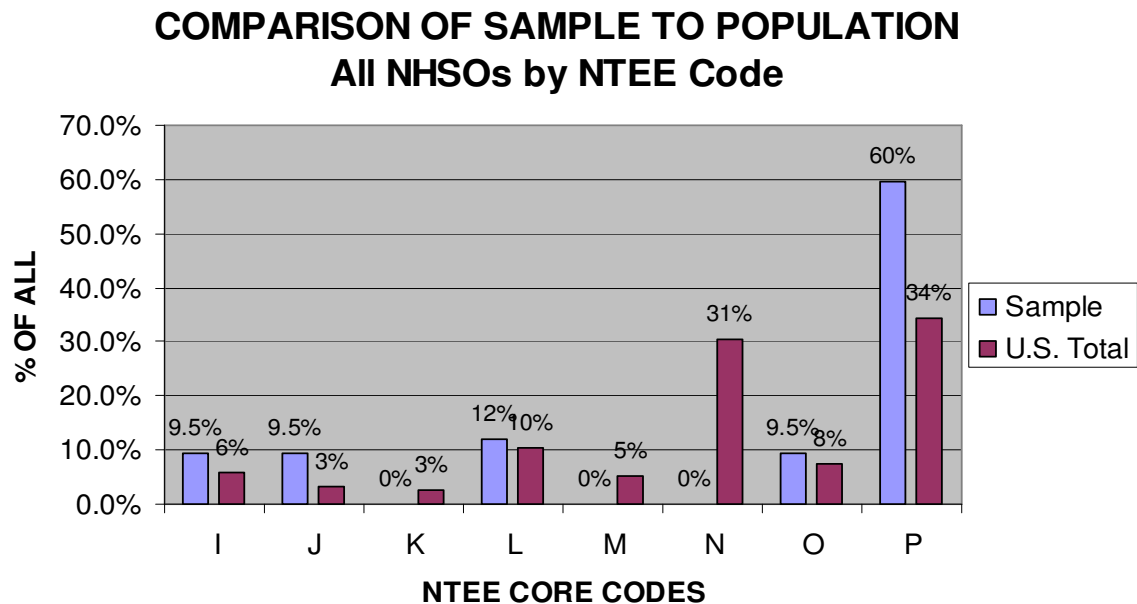
Figure 7 – Comparison of Sample to Population by County (Indianapolis MSA)



The sub categories for the Human Services sub sector include “Agriculture, Food and Nutrition; Crime and Legal Related; Employment and Occupations; General Human Services; Housing; Public Safety, Disaster Preparedness and Relief; Recreation and Sports; and Youth Development.” Figure 8 below shows how the 42 organizations

compare to the U.S. population of NHSOs by their National Taxonomy of Exempt Entities (NTEE) codes (NCCS, 2007).

Figure 8 – Comparison of Sample to Population by County – All NHSOs by NTEE Code



***NTEE CORE CODES KEY**

I=	Crime, Legal Related	M=	Public Safety
J=	Employment & Occupations	N=	Recreation, Sports, Leisure, Athletics
K=	Food, Agriculture, and Nutrition	O=	Youth Development
L=	Housing, Shelter	P=	Human Services - Multipurpose and Other

Potential interview participants were sent an introductory letters on university stationery in the mail (see Appendix A) or via e-mail that were customized for their stakeholder group and explained the nature of the study, its parameters, and that their organization was randomly selected. The letter explained that they were selected for an in-person interview that would take approximately one hour. It also explained that the interview would be audio taped for accuracy, but their individual identity and the identity of their organization would remain confidential according to the guidelines of and

approval from the Indiana University Internal Review Board's (IRB) regulations. The letter to the EDs, which were the first stakeholder group interviewed, also requested that organizational documents such as annual reports, program brochures, evaluation reports, etc. be provided for review. Stakeholders were then called approximately a week to ten days after the letter was sent, if they weren't heard from yet, to set up the interview.

Three attempts, including the initial follow-up phone call were made to potential participants. All participants were promised a copy of the research report for their respective stakeholder group. Once the stakeholders agreed to participate and selected a location, date and time of their choosing, they were sent confirmation letters in the mail (see Appendix B) referencing these details. All participants were sent thank you cards.

Sampling Procedures by Stakeholder Group

The sample procedures for each population differed according to the available information for that population and their specific relevance to this topic. After the initial forty-two organizations had been selected through for the ED sample, twenty board chairs from those same organizations were randomly selected with ten from each MSA. Introductory letters or e-mails were sent to the EDs explaining their board chairs had been selected to participate in the same research study along with a request for their board chair's contact information. This was done to acquire the most accurate and current contact information but also was intended to be a courtesy to the EDs who may be wary of a researcher speaking to their board chair. It took eleven inquiries in each MSA to acquire ten board chairs willing to participate. In the Atlanta MSA, one ED never responded with their chair's contact information and in the Indy MSA, one ED did not want to bother his board chair with this study. Once the board chair's contact

information was acquired, the same protocol was followed for them as with the EDs including introductory letters, follow up phone calls, confirmation letters, promised copies of their group's report and thank you cards.

The population database for program staff was developed from the staff roster information of the 42 participating NHSOs. This information was compiled from the organization's web sites, annual reports, program brochures or from phone calls to organizations that didn't have such information readily available. Ten from each MSA were randomly chosen to total twenty program staff in the study. Multiple program staff from some organizations, especially the larger ones were included to ensure a large enough database to procure ten willing participants from each MSA. However organizations were only represented once by program staff, as additional program staff randomly selected from an organization already represented were tossed out and only reinstated if the initial program staff person was unable to participate.

The six types of funders included in this study are community foundations, corporate foundations, independent foundations, corporate giving programs, United Ways and government agencies. The total of twenty-four funders in the study comprised two from each of the six types resulting in twelve in each MSA. The majority of the twenty-four funders were randomly selected with the exception of the major community foundation, United Way and independent foundation in each MSA. They were specifically targeted to participate because they were dominant funders of NHSOs, resulting in purposive funder samples. The Foundation Center's database was the primary source used to create the funder populations in each MSA to draw from. Funders were categorized by each type in their database and were checked to ensure they funded

NHSOs. Additional resources that were used to add funders not listed in the Foundation Center's database included the *Atlanta Business Chronicle*, the Indiana Grantmakers Alliance and GuideStar. The criterion for selecting funders was they had to have at least a minimum annual giving of \$50,000 from the most recent fiscal year. Determining the right person within the funding organization to participate in the interview was a minor issue, as for United Ways it would typically be fund distribution personnel while for foundations it typically would be program officers. In cases where there may be multiple staff with these positions within the sample organizations, only one person was randomly selected to participate.

Twenty evaluators were randomly selected with ten from each MSA. The membership lists from each area affiliates, the Atlanta Area Evaluation Association and the Indiana Evaluation Association served as the primary sources to create the evaluator population databases. They were cross-matched with the American Evaluation Association member list accessible from the members-only portion of their web site to add any not belonging to the local affiliates. The criterion for selecting evaluators was they had to have evaluated an NHSO program within the past calendar year.

Instruments

Social research is a systematic process for gathering and analyzing empirical data for the purpose of understanding the social forces and patterns at work in a given social situation (Babbie, 1995; Monette, Sullivan & DeJong, 2002; Rubin & Rubin, 2005;). This study is congruent with social research's ultimate purpose of clarifying the essence of social life (Babbie, 1995) because program evaluation is a process conducted by people in a social context, and the capacity for it relies on the actions of social actors.

One type of social research is *qualitative research* which is defined as a method of analyzing data that is not easily counted and usually comes in the form of descriptions, words, and narratives (Babbie, 1995; Rubin & Rubin, 2005). Lofland offers nine areas of focus that drives the practice of qualitative research he calls “thinking units,” and they include *meanings, practices, episodes, encounters, roles, relationships, groups, organizations* and *settlements* (societies of a small scale) (Babbie, 1995, p. 281).

Qualitative interviews are one of several techniques employed for social research including surveys, experiments, observations and analysis of existing data; and they are effective for analyzing social and political processes (Rubin & Rubin, 2005). They have been chosen for this research because: 1) the participants, whose responses were confidential, could consider what was asked of them as sensitive information; 2) evaluation capacity is greatly influenced by an organization’s environment, culture and context; and 3) the world of nonprofit human services is based on interpersonal communication and interaction between people. Each of the reasons are now discussed and substantiated separately.

Participants in this research study were informed that their responses would be confidential before they agreed to be interviewed, in their confirmation letter, and during the introduction the day of the interview. Also in the interview introduction, they were reminded that they could conclude the interview at any time or refuse to answer any questions. Despite this reassurance, the information sought by the interview instruments had the potential to be considered sensitive by the participants resulting in their anxiety or hesitation with responses. For example, an ED might internalize and personalize the interview as a reflection of the effectiveness of their leadership. A funder might consider

the intimate mechanics and dynamics of their organization's funding and decision-making processes private. Also, when asked about program evaluation regarding their grantees, the funder might also feel uncomfortable responding if their organization has not endeavored in evaluating itself to be congruent with what is asked of their grantees. Evaluators might consider questions about how they're hired and how much they're compensated to be personal matters.

Board chairs might feel uncomfortable commenting on their ED. Lastly, program staff might hesitate to respond to questions about their organization's support for program evaluation that might reflect on the organization's leadership or expose their indifference to having the program they deliver evaluated. All 126 interview participants selected the location, date and time for their interview. The majority of the interviews took place in their offices, often with their doors closed and their phones forwarded to voice mail or their assistants. A familiar, safe and comfortable environment proved helpful for the participants providing open and candid responses. The in-person, one-on-one interaction of the interview, with follow up questions and probes, enabled the interviewer to provide space for the participants to elaborate, explain and clarify their responses.

Evaluation capacity is greatly influenced by a NHSO's environment, culture and context. These interviews were conducted as a snapshot in time within these factors in order to enhance the probability that their responses would be influenced by them. These factors and their relationship to evaluation capacity could not be captured in a yes/no format or within answers that could be easily counted. Participants needed the opportunity to first be aware or reminded of these factors so they could elaborate on their responses. Follow up questions and probes conducted by the interviewer helped

stimulate this awareness or recollection, provide participants with the time to reflect and think more comprehensively about these factors, and helped the interviewer acquire the rich data necessary to describe the influence of these factors.

Human services are based upon the interaction and interpersonal communication between human beings. The primary example is NHSO program staff providing an intervention into the life of the program's consumer. These staff persons seemingly have chosen their profession based on a desire to help people, so an in-person, one-on-one interview would provide the environment and context in which they would most likely be familiar and communicate openly. The relationships between each of the participating stakeholder groups are also driven by the interaction and interpersonal communication between them which may occur in various mixes such as ED-program staff, ED-board of directors, Board-funder, funder-evaluator, evaluator-ED relationships, etc. The format of these research interviews is intended to represent that interaction and tap into the foundation of relationships that drives this world of human services.

Qualitative interviews are suitable research methods for extracting nuance and subtlety, determining the past's impact on present conditions and situations, providing a new or enhanced view on a research problem, and satisfying the need for layers of questions to acquire comprehensive information from participants on a given topic (Rubin & Rubin, 2005). For the context of this research, the nuance and subtlety might comprise how stakeholders feel about program evaluation and how those feelings might impact their actions for it. One-on-one, in-person qualitative interviews provide the added advantage of observing visual cues, in addition to their verbal responses, such as body language, facial expressions and gestures (Tripodi, 1983). The past's impact on

present conditions can come in many forms including stakeholders and their organizations' past experiences with program evaluation, how evaluation capacity has evolved in NHSOs, why funders have changed their philosophy and grantee requirements for program evaluation, etc. Providing a new perspective on evaluation capacity is one of the primary purposes for this research and can contribute to the sparse literature on the topic that only began in the late 90s. Lastly, qualitative interviews provide the forum for utilizing series or layers of questions to extract comprehensive responses necessary to describe issues of depth and breadth related to context, culture, environments, etc. such as describing the current environment for funders' support of program evaluation, describing the current state of a NHSO in terms of its financial health, mission, strategy, values and how they relate to each other.

In conclusion, the other choices of social research methods such as surveys, analysis of existing data, observations and experiments would not have been conducive for capturing the breadth and richness of data necessary to properly examine the determining factors for evaluation capacity. They would not have as effectively incorporated the environments, cultures, contexts and relationships necessary for the inquiry of this topic. For example, surveys would have allowed for more homogenous responses from every participant, while qualitative interviews provide unique conversations based on what participants recall and are willing to share. Qualitative interviews allowed participants to respond openly, candidly, and expansively while conveying their opinions, perspectives, challenges and successes. The fact that none of the 126 participants, explicitly or with the knowledge of the interviewer, refused to

answer any question or terminated any interview in progress affirms this choice as the social research method for this study.

The instruments used for the one-on-one, in-person interviews in this research were semi-structured questionnaires each designed for that specific stakeholder group (see Appendices C-G). Some questions were homogenous among all five questionnaires, such as the first question which asks for their meaning of the term of program evaluation, to enable analysis across all participants. The majority of the remainder of the questions pertain specifically to EDs (Appendix A), Board chairs (Appendix B), funders (Appendix C), evaluators (Appendix D), or program staff (Appendix E). These questionnaires were designed to be mixes of topical and cultural interviews in order to acquire information on the circumstances around program evaluation capacity as well as cultural implications such as what people deem important, pass on to other stakeholders, have learned, and what takes place in between these stakeholders in certain settings. Rubin & Rubin would categorize the breadth and subject of focus of these interviews as “In-Between,” as opposed to narrowly or broadly focused (2005, p. 5). In this category they describe the interview process as one that incorporates oral history as well as organizational culture. Specifically for this research it attempts to determine past events between these stakeholders and in their organizations as well as the lessons learned or metaphors for behavior involving activity related to program evaluation in NHSOs.

Qualitative interviews and normal, every day conversations have some common features such as questions and answers following a pattern where what someone says is determined by what the other person previously said. Interview participants are considered “*conversational partners*,” described by Rubin and Rubin as participants who

are unique individuals with their own wealth of knowledge and social constructions of reality that contribute to the flow of the interview and the direction of the research (Rubin & Rubin, 2005, p. 14). They "...may take control of the interview and change the subject, guide the tempo, or indicate that the interviewer was asking the wrong question"...or simply "...become hostile, overly friendly, threatening, or flirtatious" (Rubin & Rubin, 2005, p. 12). These situations require the interviewer to exercise quick judgment in deciding how to delicately and politely keep the interview on track and on topic, or redirect or terminate the interview.

The behavioral dynamics of the interview process have been the subject of recent discussion among survey research methodology experts. The demeanor and behavior of an interviewer including their tone of voice, the speed at which they ask questions, the amount of time allowed for participants' pausing or reflecting before answering, the balance of the interviewer talking with listening and even the interviewer's body language all guide the quality of information exchange and interaction with the participant, especially with regard to the participant's inference of what their role is in the process (Olson & Peytchev, 2007; Rubin & Rubin, 2005). Some survey research methodology experts argue that an interviewer should take a strict homogenous approach, sometimes called *formal interviewing*, by reading questions exactly as they are written and exactly in the same order to maintain the integrity and consistency of the interview, and most importantly minimize interviewer effects (Conrad & Blair in Beatty & Willis, 2007). Other experts feel that there can be advantages to taking a more informal approach by shaping the interview to be more like a conversation where the participant can feel comfortable in responding (Reamer, 1998; Rubin & Rubin, 2005).

Rubin & Rubin adamantly state that “Asking everyone the same questions makes little sense in qualitative interviewing. An interview is a window on a time and a social world that is experienced one person at a time, one incident at a time” (2005, p. 14). Recommended practices for qualitative interviewers are still evolving and more research studies on interviewing protocols and techniques are needed (Beatty, 2004; Beatty & Willis, 2007). Paul Beatty at the National Center for Health Statistics conducted research on cognitive interviewing and concluded that “...some of the most useful findings from cognitive interviews are due to adaptive, investigative skills of individual interviewers, who improvise probing based on specific interview content” (Beatty, 2004, p. 65). Beatty also indicated that while this notion might not be accepted by all researchers, such adaptive skills have implications for the training of interviewers as well as the continuing development of qualitative methodology. Some of these experts provide a third alternative for determining how best to implement a qualitative interview which is a blend of the formal and informal approaches that is more situational and based on the interview setting, the population, and other factors (Reamer, 1998).

The interview process for this research has employed this alternative mix of approaches by attempting to ask the questions as worded on each questionnaire while incorporating some probes for all stakeholder groups depending on the question, and specific probes for questions pertaining to a particular stakeholder group. There was also an attempt to ask the questions for each group in the same sequence, however quite often participants answered questions before being asked them and/or answer questions later in the interview when they have had time to think about them. Even if already answered the interviewer still asked the questions in the same sequence to follow the protocol for

consistency and provide the opportunity for the participant to expand on their previous answer. Follow-up questions typically were driven by the participant's knowledge of the question content, incomplete responses, gaps, omissions, or a new direction for an important event or context related to the research topic not previously discussed by the participant. Rubin and Rubin remind us that,

“To achieve richness and depth of understanding, those engaged in qualitative interviews listen for and then explore key words, ideas and themes using follow-up questions to encourage the interviewee to expand on what he or she has said that the researcher feels is important to the research” (2005, p. 13).

This was the basis for choosing a semi-structured questionnaire to accommodate this balanced approach to interviewing.

This choice has been affirmed, as quite often participants were able to more thoroughly respond when probed, asked follow-up questions or had questions clarified. The interviewer has observed several factors that have validated the flexibility of the interview to resemble more of a conversation. First, a transition to a more relaxed, open and candid interview resulted from the interviewer conveying he had been a nonprofit practitioner in the capacity of some of these stakeholder groups, and had interacted with all of them at some time during the past ten years. This helped break down the stereotypes and barriers often found between practitioners and academic researchers. Once participants realized the interviewer was fairly knowledgeable about their organization, their service delivery or more importantly the challenges they face by expressing empathy, they exhibited body language indicating they became more comfortable with the interview and they answered questions previously unanswered in the interview. Second, the conversation format yielded valuable information through

expanded responses to questions, resulting in participants sharing what previously they considered confidential or personal, and stimulating their interest in this research. About one third the participants took notes during the interview to capture ideas. Some said, “I didn’t think of that,” or “That’s a good idea, I think I’ll try that,” and in some cases they conveyed they had learned some things from the interview. The majority of participants requested a copy of the research report for their stakeholder group before the interviewer offered it at the conclusion of the interview.

Questionnaires were developed from larger master lists and then were prioritized and reduced in order to produce an interview of approximately 20 questions and one that would take approximately an hour. The researcher determined from past experience and from the survey research methodology research that if interviews took more than one hour, the number of participants agreeing to participate would most likely dramatically drop. *Cognitive pre-testing* is the process of administering the drafts of to a representative sample of participants for the purpose of examining their responses, gathering information about their responses, and evaluating the effectiveness of the questions in providing the information desired by the researcher. It is an important process for dealing with issues around Tourangeau’s cognitive process for answering questions which includes the sequence of *comprehension of the question; retrieval of information; judgment and estimation; and reporting an answer* (Tourangeau in Groves et al., 2004, p. 202). Groves reminds us of some of the critical areas that this process should account for such as “failure to encode the information sought; misinterpretation of the questions; forgetting and other memory problems; flawed judgment or estimation strategies; problems in formatting an answer; more or less deliberate misreporting; and

failure to follow instructions” (Groves et al., 2004, p. 209). Some of the information gathered by the cognitive pre-testing process may include:

“(1) respondent elaborations regarding how they constructed their answers, (2) explanations of what they interpret the questions to mean, (3) reports of any difficulties they had answering, or (4) anything else that sheds light on the broader circumstances that their answers were based upon” (Beatty & Willis, 2007, p. 288).

All five questionnaires were pre-tested by small groups representing each corresponding stakeholder group. Pre-test participants were interviewed using the draft questionnaires and then were asked some questions when the interviews were completed. These sessions were audio taped so the researcher could accurately review their responses and recommendations. The researcher also took notes during the process that helped with follow-up questions as well as served as the basis to improve each remaining pre-test session. Food and soft drinks were provided to each volunteer participant and parking when applicable was paid for by the researcher. The participants helped improve the questionnaires by conveying which questions they didn’t understand. They also confirmed that the one-hour limit to complete the interview was desirable.

Revisions were made to each questionnaire based on the results of these pre-test sessions. Some of the revisions included adding more probes and changing the wording of questions to increase comprehension. Additional concerns expressed by pre-test participants included those concerning *language*, *sequencing*, *use of probes*, and what they felt might be deemed *sensitive information* by other participants. The researcher attempted to balance including language that was specific to the topic of program evaluation while also using language participants could generally understand. In some cases the issue of the participant not knowing what an item was did not determine a

change in language, as not knowing was acceptable and part of the point behind the research. One example of this is the question, “Does your program have a logic model?” The majority of the Atlanta MSA EDs (62%) stated their programs had logic models, while 33% did not and 5% did not know. The majority of the EDs in the Indianapolis MSA (71%), in contrast, responded by indicating they didn’t know what a logic model was and asked the interviewer to explain it. This lack of knowledge was not a language issue but rather an indication of their exposure to an important issue for this research. Another acceptable example of a participant not knowing an answer was in responding to “How much did the evaluation cost?” The majority of EDs in indicated they did not know or did not feel comfortable providing an estimate. This lack of knowledge again is an important indicator for budgeting and other decisions made around program evaluation.

With regards to sequence, the researcher determined that the first question, “What comes to mind when you hear the term program evaluation?” would get right at the topic being investigated by the research. This first question also helps get at the heart of the cognitive process for the interview by establishing the context for the topic. The intention here was then to have them proceed with the remainder of the interview basing their responses on their personal, individual understanding of program evaluation. This follows the logic of Tourangeau’s cognitive process for answering questions. The first question in and of itself could be characterized as a *cognitive probe* in that it such probes are “...used to understand interpretation of terms, computational processes, information that was considered when answering, and level of difficulty encountered while thinking of an answer...”(Beatty, 2004, p. 58). Profile questions for each stakeholder group such

as their education, tenure in their position and others were asked at the end of the questionnaire, as they were deemed important but separate from the core content.

Skip patterns, or instructions to move to another question based on the response to a previous question, in this context were instructions for the interviewer. For example, if an ED answered question 2 indicating their organization had not made an effort to evaluate their programs, then the interviewer skipped to questions 7 through 9. Similarly, if an ED responded to question 2 affirmative having evaluated their programs, then the interviewer proceeded with questions 3 through 10. These instructions are explicitly contained in each questionnaire in upper case letters to ensure they are adhered to.

A reality of most open-ended interviews is that the order of the questions as written in the questionnaire may not exactly match the order of participants' responses, as they may answer multiple questions or individual questions at various times during the course of the conversation. For example, in describing what had been done with respect to evaluating programs in probe A of question 2, a participant may answer probes B and C by describing methods used and who conducted the evaluator. If the participant did not follow this pattern, probes were employed to attempt to get that information. If questions were answered before they were asked by the interviewer, the interviewer made notes and adhered to the protocol by asking those questions even though they were already answered. This adherence ensured consistent administering of the interviews and also reaped the occasional benefit of participants answering those questions more comprehensively than before.

Groves reminds us that "The standard for a 'good question' ...is one that minimizes the need for interviewer probing in order to get an adequate answer" (Groves

et al., 2004, p. 282). However in the context of interviewing these stakeholders, the researcher determined that establishing concrete probes in the questionnaire would enhance consistency, accommodate the fact that each participant may have their own personal, unique response, and serve the purpose of having participants elaborate on broader contextual issues. Groves also reminds us from a study by Mangione, Fowler and Louis that "...questions answered in the narrative form, which are particularly likely to require interviewer probing ...were more likely than average to be subject to interviewer effects" (Groves et al., 2004, p. 281). The researcher felt that having concrete, standard probes for each question would help reduce interviewer effects by minimizing the potential for the interviewer's inconsistent methods for following up on initial responses to main questions. Some of these probes were *anticipated probes* created for the purpose of stimulating elaboration for the more complex issues generated by the main questions while others were *conditional probes* driven by the participants' responses (Willis in Beatty & Willis, 2007, p.300).

The interviewer, however also employed additional *spontaneous* and *emergent probes* when necessary during the course of the interview (Willis in Beatty & Willis, 2007, p. 300) at the risk of inducing interviewer effect. For example, at times when the interviewer felt the participant's response was not confident or clear, a *confirmatory probe* was used to make sure the response was accurate (what the participant wanted it to be). *Expansive probes* were used when participants either did not answer completely or in cases when they began to provide valuable information and stopped. *Functional remarks*, or redirecting the participant back to the original question (Beatty in Presser et al., 2004, p. 58) were sometimes required when the participant did not answer the

question, answered another question instead or rambled on a tangent not relating to the question. In the cases of rambling, at times participants caught themselves, stopped and asked the interviewer what the original question was. In other cases the interviewer had to delicately interrupt the rambling, bring the discussion back to the issues contained in the interview, and repeat the last question asked.

Case Studies

In addition to the 126 interviews, two cases studies of NHSOs and their efforts for evaluation capacity building were also conducted. These case studies were purposive situational analyses that were intended to demonstrate examples of ideal efforts that NHSOs could use as models for considerable replication. They comprised interviews with key stakeholders, a review of organizational documentation, and the researcher's attendance at evaluation-related committee meetings. Some questions were homogenous across all interviews while others were specific to the work or historical place of that stakeholder. The review of organizational documentation included strategic plans, evaluation reports, annual reports, journal articles, board and staff meeting minutes and other examples subject to availability. Lastly, I participated in select meetings where program evaluation was discussed to observe the process, the stakeholder interaction, learn more about their process and ask clarifying questions. Notes were taken at these meetings which are also a part of these case studies.

These two organizations were chosen for several reasons. First was the convenience of their location in the Atlanta MSA and second was the disparity of their sizes with one having an annual budget of approximately \$500,000 and the other with an annual budget of approximately \$13,000,000. The most important reason was the

researcher's belief that they represent models of recommended practices with regard to evaluation capacity building.

Data Collection

All interviews with the five stakeholder groups and the case study participants were conducted in-person. They were also audio recorded and transcribed word-for-word to capture exactly what the participants were saying and not impressions of what they were saying and not through filtered through process of a note taker. Documents about the participating NHSOs were reviewed and notes were entered into a Word document for what was discovered from each in relation to the topic of ECB. The interviews and case studies took place over a three year period from 2005-2008.

Data Analysis

The 126 interviews totaled over 100 hours of tape and over 2,500 pages of transcripts. The transcripts were imported into NVIVO7®¹³ qualitative analysis software to uncover common threads, terms, concepts and themes and differences among each stakeholder group as well as for all stakeholders. The software was also used to code the data so these characteristics could be efficiently organized to serve as the basis for reporting results and developing conclusions. Pre-determined categories were intentionally not used so that the data would “do the talking” and represent itself in lieu of having to fit such categories. This represented more of an open inquiry style of research rather than having testing for a pre-determined hypothesis. The researcher determined that this approach was more appropriate for this exploratory, qualitative study that was intended to expand the understanding of and promote further research for a topic not widely studied.

The results of the data analysis were used to gain insight into the dynamics of the roles of and relationships between these stakeholder groups regarding program evaluation, and to establish conclusions and recommendations for practice. There were several methods for using NVIVO to accomplish this goal. First, the text from the responses for each question within each stakeholder group was grouped together in nodes, so they could be easily analyzed as groups of responses. At times the participants did not answer the questions in order, and for example may have answered an earlier question later in the interview due to recollection, the researcher's probes or subconsciously answering through conversation. For these interviews, the researcher searched the interview transcripts for those responses and added their text to the node to complete all of the stakeholders' responses for that particular question. Next, the researcher analyzed each node and coded them according to their common themes driven by a text search for word frequencies within the node. These became the sub nodes for which the responses would be categorized into. The remaining responses not using those specific terms placed in appropriate categories based on the researcher's interpretation of the response content.

An example that illustrates this process is the question posed to EDs about how they saw their role in the overall evaluation process. An analysis of the content of the 41 responses yielded the two most common terms of "driver" and "overseer." These became the sub nodes to segregate the 41 EDs into each category. The researcher then went back and interpreted the remaining responses not using those specific terms and placed them in appropriate categories. One example of an ED describing herself as a "driver" without using that specific term is one who stated, "I make sure it gets done." Another example

of was an ED who was personally driving the process who said, “Well, as the director of a small non-profit my job is multi-fold because I’m also the Clinical Director. I also still see clients here so I design all the evaluations. I administer one of the experiential programs that we do, so I’m responsible for administering the evaluations for those as well as those for my own clients.” Examples of “overseers” not using that specific term included an ED who said, “I am the reviewer of the information,” and one who stated, “I serve as the advisor to the board and make sure they get the information.” This process produced percentages of EDs in each category, in this case 37% drivers and 63% overseers. This information was used to determine the relationship between these roles and other important points of analysis such as whether their organizations they lead were Type I or Type II organizations.

This process was also used but expanded to cross analyze responses for questions asked of all 126 participants such as for when they were asked for what the term “program evaluation” meant to them. The large node had all 126 responses which contained the coded nodes from each stakeholder group to facilitate more effective and efficient comparisons across all groups. The results section beginning on page 199 conveys the variance for the understanding of program evaluation across the five groups which is one of the most critical findings of this study supporting a an equally important conclusion.

Limitations of the Study

Semi-structured, one-on-one interviews using non-directive probes were used as the primary method of information acquisition for this study. The utilization of a semi-structured questionnaire typically presents opportunities for bias, variance and other

research limitations. A review of interview transcripts confirms that some variance exists among the interviews, as follow up questions were driven by the specific responses of the individuals being interviewed. The interviewer concluded that at times an interview had to be more of a conversation in order to acquire the necessary information from the participant, therefore validating some flexibility in these interviews. However the view as to whether this is advantageous or it violates recommended practices is left up to the survey research methodology experts.

Qualitative interviews inherently present the opportunity for interview effects including interviewer bias. The interviewer's interest and excitement over this topic occasionally showed through during an interview and breached the protocol rule for intended neutrality that does not permit expressing personal opinions or contributing to the pull of the conversation (Holstein & Gubrium, 1995). Interviews have also been characterized as a means for producing meaning, as the interaction between an interviewer and the respondent contributes to the social construction of reality (Berger & Luckman, 1967; Holstein & Gubrium, 1995). This can be viewed as both a limitation and an advantage depending on one's own perspective. Interviews were chosen over written surveys simply because surveys would not have yielded the level of candidness or expansiveness in open responses that were received from the participants. The interviewer also concluded that the interview process would produce the meanings and results that would properly address the issues related to this research topic. This perspective supports the notion that interview responses are natural, practical productions of such interaction that would be a part of any process that socially constructs meaning (Holstein & Gubrium, 1995).

Undercoverage occurs when “...eligible members of the population cannot appear in any sample drawn for the survey” (Groves et al., 2004, p. 70). GuideStar was used as the source to develop the database of the population from which the samples would be drawn. The Atlanta MSA NISO population comprised 587 organizations while the Indianapolis MSA population had 239. “GuideStar obtains information from the IRS Business Master File, IRS Forms 990, 990-EZ, and 990-PF, and individual organizations” (GuideStar, 2005). As a result there may be organizations that fit the income level or geographic stratification criteria of the sample frame but are not listed on GuideStar to be eligible to be drawn for the sample. Some possible reasons for this include new organizations who had not yet filed their 990 at the time the samples were drawn, organizations who had not filed their 990 for the current fiscal year at the time the samples were drawn due to being granted an extension, or organizations that had filed but GuideStar had yet received the 990 from the IRS or uploaded it to their web site. The tradeoff in using GuideStar, even with its minor imperfections, is that the organization offers the most easily accessible and cost effective searchable database of nonprofit organizations.

Undercoverage may occur in the selection process of the funders and evaluators. It is possible, but unlikely, that some funders were left out of the sampling process if they were not listed in the sources used to build the population database. Similarly with evaluators, it might be possible that some were left out because they were not members of their local affiliate association or the national organization, the American Evaluation Association. Overall, the researcher’s contention is that these instances are in the extreme minority, but it is important to note for the ethical reporting of research

limitations that the selection process was not perfectly complete with respect to coverage. The sample of 42 participating NHSOs was not stratified according to their NTEE codes within the human services subsector. As a result, undercoverage existed for how the forty-two organizations compared to the breakout of NTEE core codes within the U.S. population of NHSOs according to NCCS (NCCS, 2007). There were no organizations in the sample representing the NTEE core codes for Food, Agriculture, and Nutrition which represented 3%; Public Safety which represented 5%; or Recreation, Sports, Leisure, Athletics which represented 31% of all NHSOs in the U.S. (see Figure 8). The sample frames used for both the Atlanta and Indianapolis MSAs reveals that six of the organizations that did not respond to participate in the study were Recreation, Sports, Leisure, Athletics organizations indicating there may be special challenges in reaching such organizations to participate in this kind of research.

A more prominent influence for selection bias is the simple fact that even with random selection the samples of participants were driven by their willingness to participate. For example, in the Atlanta MSA portion of the study, 52 organizations were randomly selected from a population of 587 before 21 EDs agreed to participate whose organizations fit the sample frame, accounting for a 40% success rate. Two EDs refused to participate, one had major surgery and the other 28 did not respond after three attempts to contact them. One of the EDs who refused to participate did not indicate his reasons while the other claimed he wouldn't reveal information that was "private and confidential" even though it was explained to him that IRS 990 forms are public information and that as a Housing and Urban Development (HUD)-assisted nonprofit organization, their annual report was also public information.

In the Indianapolis MSA sample it took 36 inquiries randomly drawn from a population of 239 organizations to yield 21 participating EDs yielding a success rate of 58%. Three declined to participate, one with no reason, another that stated she was “too busy running a homeless shelter and having staffing problems” and the third indicated her organization was having an annual audit and she was in the middle of two big projects. One phone number had been disconnected with no forwarding information while another was continuously being checked for trouble. One ED had just retired, and the organization had not hired her replacement yet. The remaining nine did not respond.

Ineligible elements are examples of sample elements that do not fit the sample frame. The key in dealing with them is to remove them before the selection process begins. This issue cropped up in using GuideStar, as their classification system is imperfect. While building the sample frame from their database, several organizations clearly belonging to other sub sector categories such as Arts, Culture and Humanities, Health, Environment and Animals, Public Benefit or others were not included in the population databases for each MSA study. GuideStar’s sub categories for the Human Services sub sector include “Agriculture, Food and Nutrition; Crime and Legal Related; Employment and Occupations; General Human Services; Housing; Public Safety, Disaster Preparedness and Relief; Recreation and Sports; and Youth Development” (GuideStar, 2007). If it was determined that an organization was erroneously categorized in human services and did not fit into any of these designated sub categories, they were not included in the population database. While this required a judgment call from the researcher, and objective outsiders might consider this a form of bias, GuideStar’s

categories made such judgments fairly clear and such instances were typically below 5% of all organizations in each of the MSA's population database building processes.

Limitations also existed in the cognitive pre-testing process. Participants were selected by the researcher through his available contacts in the appropriate stakeholder groups for reasons of time and to maximize the chance for their voluntary participation. By survey research methodology recommended practices, the pre-test groups were relatively small, ranging from two to eight participants depending on the group. The pre-test process was rather basic and did not include a more comprehensive approach referenced by survey research methodology experts due to the typical compromises for time and cost. One example that would be included in such a pre-testing process would include comparing pre-test methods for efficiency and effectiveness, and/or the intended incorporation of more than one. Another example would be incorporating behavior coding which helps identify problems in the interview process, based on the interviewer's or respondent's behavior, related to issues of comprehension of the questions, and the mapping of their judgment for the format of their responses (Holbrook, Cho & Johnson, 2006).

The same aforementioned limitations for the interviews used for the five stakeholder groups would apply for the interviews used in the two case studies. However, organizational documents in the case studies were more accessible due to the organizations' willing participation and cooperation. This resulted in a more thorough review of documents, which would be expected for case studies, that was used as a source of information to discover congruence with the participants' responses and

provide a way to corroborate evidence. Such documents at times were beneficial in that they provided detail and specificity not conveyed in the stakeholder interviews.

Selection bias was inherent in the purposive selection of the two organizations participating in the case studies, Families First and Decatur Cooperative Ministry. Pure case study research would call for a random selection and a neutral analysis of the unit(s) of observation. However, these case studies were purposely chosen to serve as ideal examples and potential models for other NHSOs to learn from concerning ECB, therefore automatically injecting subjective bias. The researcher determined that it was important to balance the more prominent challenges found in the 126 interviews with some success stories. Depending on one's perspective this can be viewed as a limitation. The geographic convenience of selecting two organizations in the metro Atlanta area also can be viewed as a limitation, in that the analysis may include environmental factors typical or exclusive for that area.

The researcher was familiar with these organizations through the network of the metro Atlanta nonprofit community, and in the case of DCM, had personally been involved with the organization through volunteering. The organizations were chosen for their exemplary efforts in building evaluation capacity, and together they demonstrated what a large NHSO and small one could do in this area. Such efforts included, but were not limited to, strong ED and board participation and support for ECB, a desire to publish evaluation results and share them with the general public, and an overall effort to embed program evaluation into the culture of their organization.

Caution is offered here for generalizing the results of this study due to the small sample sizes used for the interviews and the case studies. First, we would not want to

generalize what all community foundations' role in evaluation capacity might be based on four included in this study or generalize what all human services program staff's role might be in evaluation capacity due to the great diversity of service delivery types and professions within the subsector. Second, studying organizations, stakeholders and the issues affecting this topic in only two MSAs in the U.S. would also preclude generalizability for not being geographically representative. Third, a time gap between stakeholder interviews exists because of time limitations of the researcher and the inability to interview 126 people within relatively the same timeframe. For example, the ED interviews took place in 2005 while their program staff were interviewed primarily in 2008. Interviews are snapshots in time, so any inference about the similarities or differences among stakeholder perspectives should take this into account.

The researcher contends that generalizability is not the issue it would be if this study utilized a statistically representative samples of the NHSOs stakeholder populations, however that would not serve well as the means for the purpose of this study as noted before in the methodology section. Qualitative interviews and the resulting qualitative analysis serve the purpose of discovering variation, uncovering slight differences in meaning, and overall examining a complex issue driven by human interaction and depict it in a way that it can be better understood (Rubin & Rubin, 2005). The primary goal of this study was not to be able to generalize data, but to inquire into what the determining factors are behind a complex issue driven by multiple stakeholders and not widely studied.

Results

The results of this study come from a total of 126 interviews conducted with five different stakeholder groups. These results are broken out into seven sections. The first section reports characteristics for all 42 NHSOs that participated in the study. The next five sections contain results specific to each of the five stakeholder groups – EDs, Board Chairs, Program staff, Funders, and Evaluators. The last section covers two case studies of NHSOs involved in ECB.

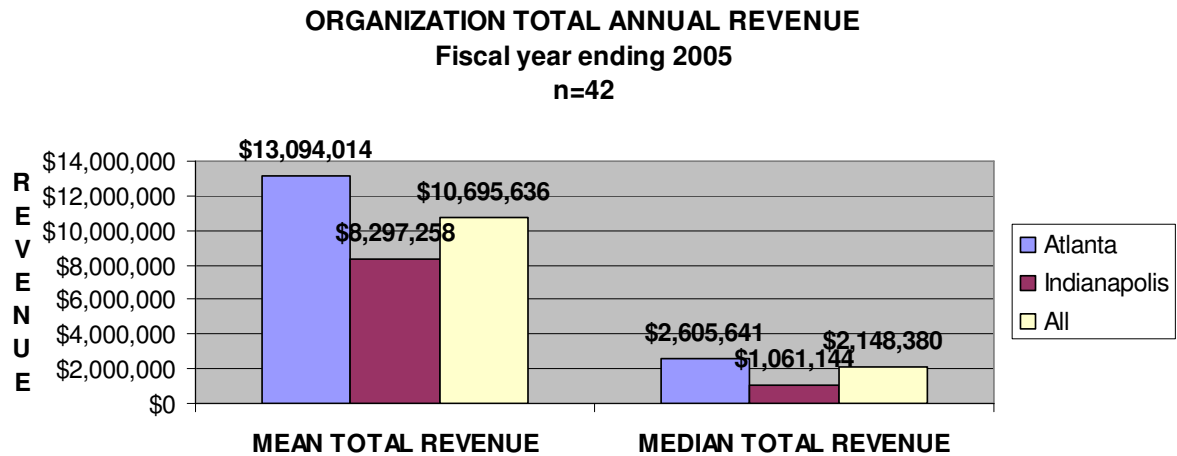
NHSOs

Approximately 17% of the organizations are faith-based while approximately 71% of the organizations have one or more government contracts with various federal, state and county agencies, such as Housing and Urban Development (HUD), Department of Human Resources (DHR), Centers for Disease Control (CDC), and others. There are a few notable changes in some of the 42 NHSOs that took place during 2005 through 2008 when this research was conducted. Nine of the organizations had changes in executive directors. Of those nine, three retired and one changed positions to lead the organization's operating foundation. One organization was absorbed into the larger metro organization for the same services, and one organization went out of business.

The total revenue for each of the 42 participating NHSOs was taken from line 12 of their IRS 990 forms for their fiscal years ending in 2005 and they ranged from \$284,634 to \$136,962,789. The organization with the most revenue was a national headquarters office serving as an outlier with more than \$100,000 greater than the organization with the next highest total revenue at \$36,854,660. The influence of this outlier in skewing the mean total revenue for all NHSOs upward can be seen below in

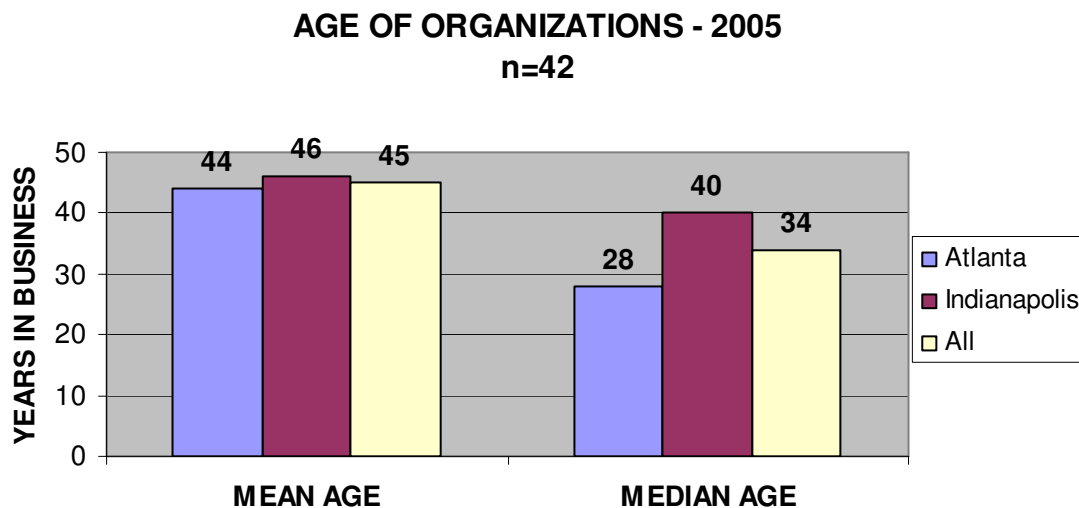
Figure 9 that displays the mean and median total revenues for the 42 participating organizations.

Figure 9 – Organization Total Annual Revenue



The organizations varied in how long they existed by 2005, ranging from 151 to 3 years. Their mean and median ages are shown below in Figure 10.

Figure 10 – Age of Organizations



There was variance in how active and comprehensive the NHSOs engaged in evaluating their programs. The data for this activity served as the basis for developing criteria and a typology of three categories of organizations as shown below in Table 6.

Table 6 – Organization Rating Types for Program Evaluation Activity

ORGANIZATION TYPE	PROGRAM EVALUATION CHARACTERISTICS
TYPE I	<ul style="list-style-type: none"> • Comprehensive and usually mixed evaluation methods reflecting a balanced approach • Program-driven and outcome-focused rather than predominantly driven by external stakeholders • Staff dedicated to evaluation and/or external evaluators conduct the evaluation • Evaluation results used primarily to improve programs • Demonstration of long-term commitment for and activity in program evaluation • Efforts to ensure the evaluation process improves and evolves
TYPE II	<ul style="list-style-type: none"> • Primarily a single evaluation method usually comprising a consumer survey • Focus on consumer satisfaction and external stakeholder requirements • Staff conduct the evaluation on an ad hoc, as needed basis • Evaluation results used primarily for program alterations relating to process i.e. logistics, location, etc. • Sporadic commitment for and activity in program evaluation limited by lacking capacity and/or driven by changes in external stakeholder demands • Evaluation process changes little or not at all
TYPE III	<ul style="list-style-type: none"> • No effort to evaluate program(s) indicated

These criteria were not pre-determined but instead based on the organizations' activity in program evaluation. This information came from the interviews with EDs, Board Chairs, and program staff, as well as a review of the organizations' documents and web sites in an attempt to assemble as accurate a picture of their evaluation efforts as possible.¹⁴ This information was then used to determine whether the organization fell into the Type I, II or III categories as shown above in Table 6.

This typology is not definitive and its purpose is to illustrate the general divisions among the characteristics of the program evaluation efforts from the participating

organizations. It is important to note that these divisions were not always exclusive, as a few organizations had a mix of characteristics between categories and arguably overlapped a bit between Type I and Type II, or were seemingly in transition between these categories. Figure 11 below shows the percentage of organizations in each category.

Figure 11 – Organization Type by Program Activity

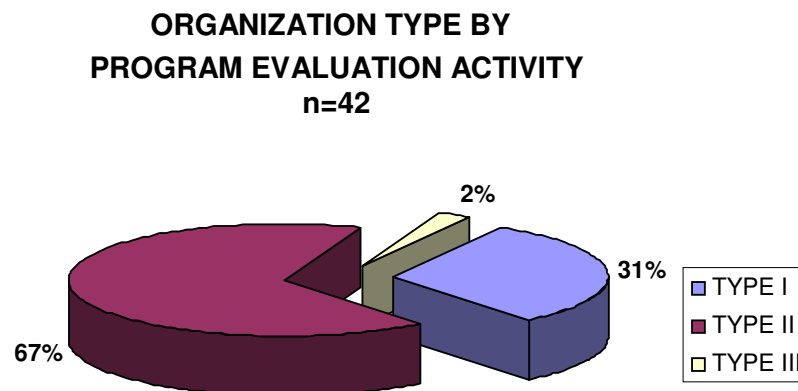


Table 7 below shows some of the organizational characteristics of the NHSOs in each of the three categories.

Table 7 – Organizational Characteristics by Type

RATING CATEGORY	ATL MSA	INDY MSA	MEAN REVENUE	MEDIAN REVENUE	MEAN AGE	MEDIAN AGE
TYPE I	54%	46%	\$24,058,001	\$12,763,376	61	67
TYPE II	46%	54%	\$4,814,546	\$920,775	38	28
TYPE III*	100%	0%	\$1,691,118	\$1,691,118	28	28

*Note: 1 organization

The total revenue of these organizations appears to have a relationship with their category type. There are several possibilities for what is driving this relationship. Larger

organizations simply have more resources, structures and systems, and might be more likely to allocate some to engage in program evaluation. Larger organizations typically have more external demands from funders, government agencies and accreditation bodies, so they are more likely to engage in program evaluation to satisfy these demands. Another possibility is that these organizations may have been the “progressive frontrunners” that engaged in program evaluation well before the surge in such activity in the 90s and therefore have benefited by being better equipped to acquire more resources from funders. Similarly, older organizations are more likely to have steadily increased their annual total revenue and have allocated the resources, developed the systems and structures necessary for program evaluation.

Executive Directors

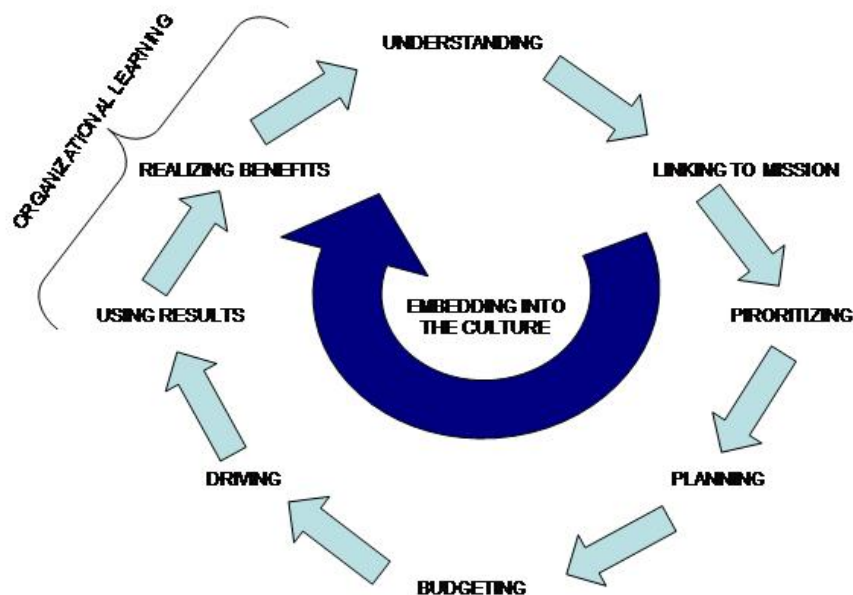
All 42 EDs were asked what comes to mind when they hear the term program evaluation in order to get their own, individual perception of it. Their responses have been split up into two corresponding categories to represent how they described program evaluation as an *activity* and their *depiction* of it based on their opinions, experiences and impressions. Some variance in perceptions existed across the 42 ED interviews, likely due to a number of factors including, but not limited to, their tenure, the size and age of their organization, their education level and previous exposure to evaluation, the level of their organization’s engagement in government contracting and accreditation, the level of influence from their board and conversely their ability to influence their board, and personal experience with the challenges for program evaluation such as time and resource constraints. In spite of this variance there were some fairly common descriptions that emerged from their responses that are displayed below in Table 8.

Table 8 – Common ED Perceptions for Program Evaluation

ACTIVITY	DEPICTION
<u>Action</u> <ul style="list-style-type: none"> • Determining • Measuring • Looking at • Assessing 	<ul style="list-style-type: none"> • Important • Compliance • Responsibility • Change • Difficult
<u>Items</u> <ul style="list-style-type: none"> • Outcomes • Consumers' needs • Program's effectiveness • Services' effectiveness 	<ul style="list-style-type: none"> • Expensive • Benchmarking • Quality assurance • Challenging • Time consuming • Informal

The 42 EDs varied in the decisions they made and actions they took for program evaluation in their organizations. Eight action steps emerged from the interviews for which some EDs engaged in all, some, or none of them. These steps are illustrated below in Figure 12:

Figure 12 – Executive Director Actions for Program Evaluation



These eight steps are portrayed as a continuous process that contributes to organizational learning, embedding evaluation into their organization's culture and ultimately for ECB. These two specific contributions were common themes from the EDs who took all eight steps. They typically referenced organizational learning at the combination of *using results* and *realizing benefits*, described as a turning point where staff attitudes, particularly program staff's, changed from anxiety or fear for evaluation to interest and excitement. These EDs described these eight steps as a continuous process and they made comments referencing or describing the embedding of program evaluation in their organization's culture. Some examples of their comments included "institutionalizing evaluation in our organization," "resulting in evaluation becoming a part of our organization's culture," "program evaluation is just part of what we do," "it has been conducted here for years and our EDs, including myself have been hired based on our experience and support for it," and "it's such a part of what we do here, I can't imagine my successor steering away from it." All of the EDs that took all of the eight action steps and a few that took the majority of them lead Type I organizations, while the majority of the other EDs that took less than these eight action steps lead Type II organizations. One ED who did not report any effort to have her organization's programs evaluated lead a Type III organization. The eight steps will now be discussed individually.

There is variance in how leaders described their *understanding* of program evaluation. The EDs that characterized program evaluation within the context of measuring the effectiveness of programs tend to lead Type I organizations. The EDs that characterized program evaluation as more of a management or administrative task

typically lead Type II organizations. Similarly, the majority of the EDs describing program evaluation as a systematic process that is an integral part of their organization's operations lead Type I organizations while those describing it exclusively as assessing consumer satisfaction or as a requirement for external stakeholder demands tended to lead Type II organizations.

The EDs were asked if program evaluation was *related to their organization's mission* in any way to determine if these issues were aligned or if program evaluation was treated as a separate activity not necessarily driven by their mission. The majority of all EDs (93%) confirmed that program evaluation was linked to their organization's mission. Some examples of their descriptions of this relationship included:

- "I think in order to be able to accomplish our mission we have to evaluate what we're doing. It is sort of like the roadmap to get us to continually fulfill our mission."
- "Yes because the fact that our mission is to put people to work, and we have to know how many people went to work and where they went to work, was the training and/or the placement service and/or the job successful in getting to that outcome, to fulfilling our mission."
- "Surveys and program evaluation are very important so we stay where our mission is."
- "Yes, it is. By looking at the evaluation of services we provide and the number of persons that we're able to provide the services to."
- "Yes, our mission is to assist people in fully participating, so measuring those things ensures that's what we're doing."

The EDs that reported cascading this value orientation about performance and alignment with their mission throughout their organization lead Type I organizations. They characterized program evaluation as a process to be integrated with other important tools and systems for performance, such as strategic planning, performance budgeting, and resource allocation. Those that described this relationship but did not use the mission as a value orientation lead Type II organizations.

The EDs' top three *priorities* fell primarily into three categories - financial stability, program effectiveness, and staffing. The most frequently identified top priority by 40% of the respondents was financial stability. The most frequent second priority identified by 31% of the respondents was program effectiveness. The most frequently reported third priority identified by 21% of the respondents concerned staffing issues such as hiring, training, and professional development. Of the 48% of the EDs who chose program effectiveness as a top three priority, 33% ranked it as their first priority, 54% as their second priority and 13% as their third priority. EDs that chose program effectiveness as one of their top two priorities tended to lead Type I organizations.

EDs were asked about whether program evaluation was a part of their strategic *planning*. Of the 34 organizations (81%) having a current strategic plan, 73% of their EDs stated that program evaluation was included in their plans. Slightly less than half of these EDs (43%) played a part, either by themselves or in conjunction with other stakeholders such as their Board of Directors, in ensuring program evaluation was included in their organization's strategic plan. These EDs typically lead Type I organizations. One ED indicated that their annual program evaluation results drove their organization's strategic plan. Their plan's goals and objectives were based on recommendations for improving the program and the outcomes for their consumers that came from the evaluation.

Approximately 60% of the EDs played a part in ensuring their organization's evaluation efforts were *budgeted*. Of the 41 organizations engaging in program evaluation, 88% of them funded their evaluation efforts solely out of their operating budgets; while 10% funded their efforts through a combination of operating budgets and

external sources such as grants, and 2% funded their efforts solely through grants.

Variance existed between how Type I and II organizations allocated, budgeted and managed financial resources for program evaluation yielding no clear trend.

EDs described how they saw their *role in the program evaluation process*. The majority of the EDs describing themselves as “drivers” of the evaluation process lead Type I organizations. Some examples of how they characterized their role include:

- “Making sure the organization participates in evaluation because it is important”
- “My role is critical to the whole process because if someone’s not driving this at the top of the organization, I have learned that it won’t happen.”
- “Making sure it gets done”
- “If I don’t insist that’s a part of the way we do things, and we will make every effort to do that, it’s not going to happen. If the executive director doesn’t believe it is a priority, I don’t think it’s going to happen.”
- “Making sure it continues to happen”
- “Deciding who’s on point for program evaluation and how much time is spent on it.”
- “Well, I think my role is really driving it to make sure that it is done in a professional manner and we can use that information.”

The EDs characterizing themselves in less involved role in the program evaluation process described themselves as “overseers” and the majority of them lead Type II organizations. Some of the aspects of their role included reviewing information, making recommendations, encouraging staff, providing direction, and helping the process get started.

EDs described how their organization *used results* of the evaluation and how they and other stakeholders *realized the benefits* of program evaluation. The majority of the EDs leading Type I organizations described these two steps together as the point where organizational learning took place. All of the several EDs that referenced program

evaluation as a learning opportunity for their organization lead Type I organizations.

They indicated that organizational learning took place when stakeholders, especially staff and the board, realized how evaluation can be used to improve the program(s), the organization's work towards their mission, and their individual work performance.

Program staff, in particular, valued learning how to improve service to their consumers.

Approximately 67% of the EDs stated their organizations use evaluation information to alter their program(s). This was the most frequent at 34% of all responses. The next two most frequent of all responses were fundraising at 17% and staffing at 11%. Of those leaders who indicated they alter their program(s), 18% specifically referenced "improving their programs" which all lead Type I organizations. Of the 41 organizations engaging in program evaluation, 100% of them stated their organization benefited from it. The majority of the EDs referenced several benefits; however 24% of them indicated that having information to demonstrate program effectiveness was their most important benefit comprising the most frequent response.

Some of the profile information gathered from the EDs is compared between Type I, Type II, and Type III organizations as shown below in Table 9.

Table 9 – ED Profile Information by Organization Type

RATING CATEGORY	MEAN TENURE	MEDIAN TENTURE	RECEIVED INSTRUCTION** IN EVAL.	HIGHEST LEVEL OF EDUCATION
TYPE I	12.8 years	10 years	100%	69% - Master's 31% - Bachelor's
TYPE II	10.4 years	8.5 years	68%	46% - Master's 46% - Bachelor's 7% - H.S. grad
TYPE III*	27 years	27 years	0%	100% - H.S. grad

*Note: 1 ED

**Instruction could be a college course, part of a college course, a workshop, a seminar, learning on the job or more than one of these examples

The EDs that have received instruction in evaluation and those who have higher education levels were more likely to lead Type I organizations. Instruction in evaluation would likely contribute to dispelling myths, gaining a better understanding, and increasing the chance the ED would engage in such activity in the future. This particularly was the case for EDs with Master's in Social Work (MSWs) who received instruction in evaluation as part of their academic experience and their chosen career path. Learning about the importance of program evaluation in this setting might have different results than if they learned it as a result of demands from external stakeholders. In the latter situation, they are more likely to view program evaluation as something they have to do in order to get the grant, government contract or accreditation as opposed to something they should link to their organization's mission and embed into their organization's culture so it can become a basic assumption.

Higher education levels would also increase the likelihood they were exposed to evaluation earlier in their academic or professional careers. Courses in program evaluation mostly reside in masters or doctoral programs. An advanced education is more likely to contribute to an ED's understanding of how program evaluation can be used as a management tool, especially for those that have taken business, management or public administration courses.

Board Chairs

The perceptions of program evaluation from the 20 board chairs varied greatly, as overall there were no common characterizations of how they described program evaluation. However, a few observations are helpful in contrasting their perceptions with other stakeholder groups. First, terms normally associated with describing program

evaluation were hardly used. Four board chairs referenced “outcomes;” three referenced “analysis;” three referenced “the effectiveness of the program;” two talked about “improving the program;” and one mentioned “measuring performance.” Second, more board chairs talked about issues centered on the resources for the program in terms of inputs and efficiency. They discussed allocating resources for the program and how resources were used in the program. Third, several board chairs seemed confused and either described the program or discussed the evaluation of the organization. Fourth, discussing the meeting of goals or objectives trumped discussion about the utilization of evaluation information which was only referenced by two board chairs. One provided her perception that most accurately described program evaluation among the 20 respondents:

“Looking at all components, and the processes, and outcomes of a program and looking at it as what works well as well as how you would improve that program.”

Only half of the board chairs provided their depictions of program evaluation with none of them stated more than once. Some included “mission centric,” “complicated,” “formal,” “rigorous,” “objective,” and “difficult.”

The board chairs were asked if program evaluation was related to their organization’s mission and the majority (95%) confirmed that program evaluation was linked to their organization’s mission. Some examples of their descriptions of this relationship include:

- “I would think that they are alpha and omega.”
- “It’s like the hand and the glove. You have to evaluate in order to know if you are fulfilling your mission. You don’t know if you have your mission covered if you don’t evaluate.”

- “Well yeah, theoretically the programs are there to fulfill the mission; right? So if the programs aren't doing that, then you're mission statement is kind of, a joke. So, you either have to revise the mission statement to reflect what you're actually doing, or if the mission statement is accurate and correct, then you got to change your programs and get them in line with what you say you're really doing.”
- “Yeah, one of the direct links is we are evaluating programs that impact our youth.”
- “I think that there's a symbiotic relationship, if you will, in the sense that the evaluations certainly ensure that you are meeting the mission that you've set forth to meet. I think secondly the evaluation process invariably will cause you to question certain things.”
- “Yes, because we're looking at the person to be able to live independently with a high level of quality. And so we have to be concerned about quality.”
- “Well, I think the evaluation, when you sum it all up, is really an evaluation of how well you're meeting your mission.”

The board chairs were also asked to describe what accountability meant for the board as a governing body and separately what it meant to their organization. A summary of their categorical responses is shown below in Table 10.

Table 10 – Board Chairs’ Descriptions of Accountability

WHAT ACCOUNTABILITY MEANS	FOR THE BOARD AS THE GOVERNING BODY	FOR THE ORGANIZATION
Fiscal accountability	26.5%	0%
Following through on things the board promises to do	14.7%	N/A
Ensuring programs are effective	8.8%	30.0%
Being accountable to stakeholders (funders, consumers, partner organizations, community, and/or each other on the board)	8.8%	30.0%
Carrying out the mission and vision of the organization	8.8%	16.7%
Holding staff (including ED) accountable for carrying out their responsibilities and meeting their objectives	8.8%	0.0%
Overseeing the governance of the organization	8.8%	N/A
Legal accountability	5.9%	0.0%
Maintaining ethical standards	2.9%	6.7%
Promote the organization	2.9%	0.0%
Be a resource for staff – expertise, guidance, etc.	2.9%	N/A
Ensuring resources are used efficiently	0.0%	16.7%

Some examples of their actual responses included:

Accountability for the Board

- “We as a board are charged with assuring that programs are well executed. We need to have projects in place that allow us to achieve that assurance.”
- “I don't know that we have a very strong accountability structure in place for our board. We're not very good right now at following through on the things we say we're going to do.”
- “Well, we are responsible for the governance of the organization, and that includes having enough funds to run the organization as well as the programs.”

Accountability for the organization

- “Communicating to our supporters, the state, and all the key stakeholders that we are good stewards of their money and that we are focused on achieving the most appropriate results we can, given the set of factors that we're dealing with.”
- “Accountability would be to our villager to provide them the best care and make sure we meet all the rules, regulations, laws, codes and so forth that govern this operation.”
- “I think as an organization we are accountable to our clients and to the greater community. Accountability means we're helping survivors and victims of sexual assault deal with what has happened to them, and become better able to adapt in society and to claim their lives again.”

A review of Table 10 indicates that a board's emphasis on the fiscal issues for a NHSO, while important and part of their responsibility may be creating an imbalance regarding programmatic responsibility. This is supported the fact that only a few board members referenced program effectiveness as a board responsibility but 30% referenced it as a responsibility for the organization, one of the two most frequent responses. This might also indicate a separation between policy and programs, where the organization's leadership, management style and culture perpetuate a divide between these two areas. Program related issues are typically categorized as “operational” or “day-to-day” and therefore do not garner the proper level of attention, discussion and responsibility from the board for ensuring that programs are attaining their intended outcomes for consumers.

The other most frequent response at 30% was accountability to stakeholders, which would seemingly contribute to program effectiveness garnering more attention from the board. However, in this their characterizations of this accountability regarded external stakeholders such as the community, funders or partnering organizations, and internal stakeholders such as the board, staff and consumers. Only a few referenced consumers which might also explain the low percentage of responses indicating program

effectiveness was their responsibility. Overall, these particular responses paint a picture that these board chairs primarily see themselves supporting each other, supporting the staff and ensuring that demands and expectations from those outside the organization are met.

This section discusses the board's participation in ECB. Issues related to their participation include what their top three priorities were; their ED's and their own description of the board's role in the program evaluation process; whether program evaluation has been discussed at board meetings; the board's role in including program evaluation in strategic planning and the organization's budget; and who on the board, if anyone, was responsible for program evaluation in their organization.

The two most frequently cited top, second and third priorities for the board and board chair, from the board chair's perspective, are displayed below in Table 11:

Table 11 – Top 3 Priorities for the Board and Board Chair

	TOP PRIORITY	SECOND PRIORITY	THIRD PRIORITY
Board	1) Financial stability (30%) 2) Achieving the mission (15%)	1) Financial stability & Board management (tied) (20%)	1) Financial stability (25%) 2) Board management & Connections with constituency (tied) (15%)
Board Chair	1) Board management (20%) 2) Financial stability (15%)	1) Board management (25%) 2) Financial stability & Operations management (tied) (10%)	1) Board management (25%) 2) Financial stability & Connections with constituency (tied) (15%)

*NOTE: 40% of the board chairs had the same top three priorities for the board and themselves

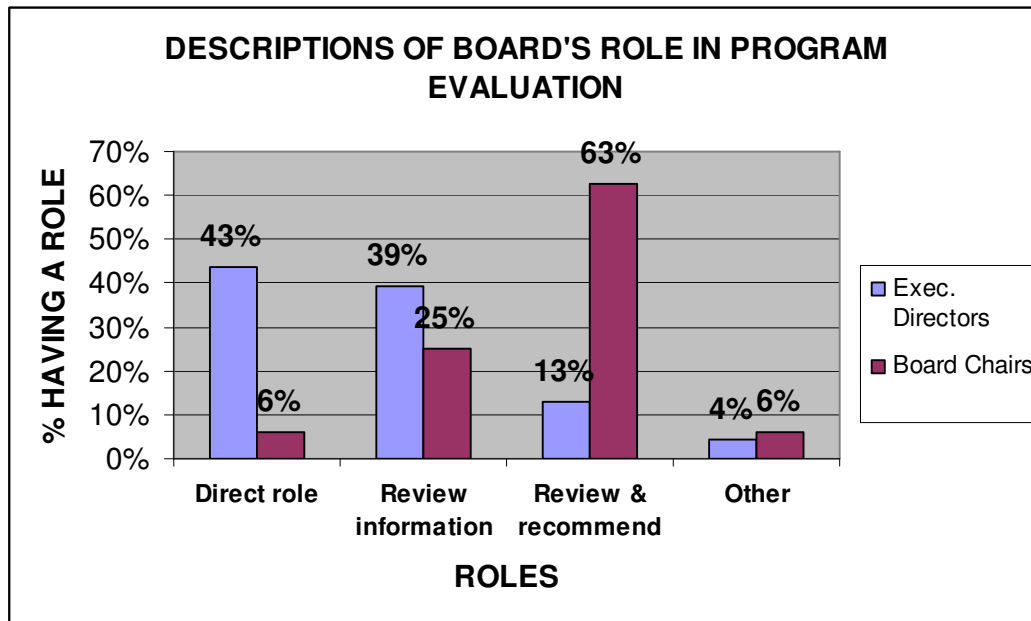
Other priorities included *Staffing*, *Program effectiveness*, *Capital improvements*, *Establishing the vision*, and *Integrity and professionalism*. The results shown in Table 11 are congruent with those in Table 10 as we would expect something not predominantly referenced as part of the board's accountability to also not be a priority. It is important to

note with regard to ECB that only one of the 20 board chairs had program effectiveness as a priority, in this case his second priority. He was the board chair of a Type I organization. The results in Table 11 also reflect the similar emphasis in Table 10 on fiscal matters.

Approximately 38% of board chairs had some role in ensuring that program evaluation was included in their organization's strategic plan, according to their EDs. Of those, half referenced the board alone had this role, slightly less than one third referenced the board with the ED, and approximately one fifth referenced the board with staff including the ED. Approximately 53% of the Type I NHSOs had their board have a role in ensuring program evaluation was in the organization's strategic plan while 39% of the Type II organizations had their board in this role. Approximately 46% of the Type 1 and 39% of the Type II NHSOs had their board playing a role in ensuring program evaluation was budgeted for, according to their EDs.

Approximately 54% of the EDs and 80% of the board chairs interviewed indicated their board had some role in the program m evaluation process. Figure 13 below shows how they each described the board's role.

Figure 13 – Descriptions of Board’s Role in Program Evaluation



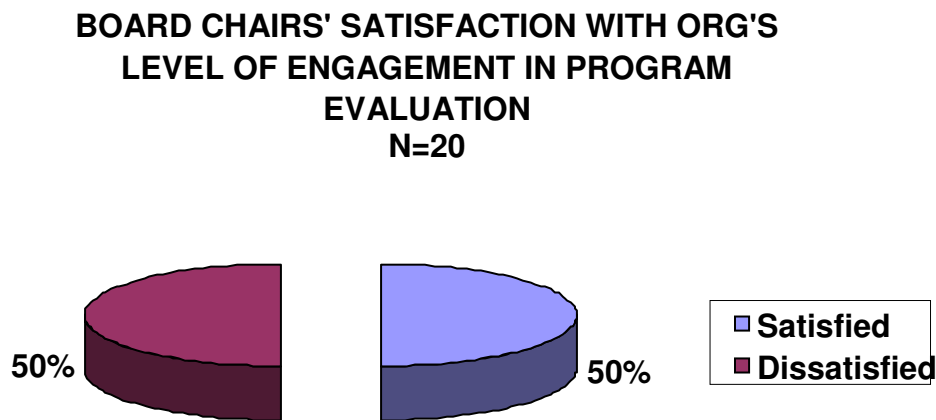
KEY: **Direct role** = board directly participates in the evaluation process
Review information = board reviews evaluation information provided through a report and/or presentation
Review & Recommend = board reviews evaluation information and makes recommendations for goals, actions or improvement
Other = “don’t know,” “funds the program,” “committed to outcomes and quality,” “oversight.”

The majority of the board chairs (85%) indicated that program evaluation has been discussed at board meetings during their tenure on the board. The frequency for when it was discussed varied from as often to three to four times a month to annually, and 25% stated it was discussed at every board meeting. These frequencies can be misleading, as about half of the board chairs when affirming it as a meeting topic described it as a process such as evaluating annual business goals, strategic plans, or the finances of the organization. Conversely, the organization discussing it three to four times a month has an evaluation committee on the board, and two of the organizations

discussing it every month have program assessment and quality assurance committees on their boards. All three organizations are Type I organizations.

Board chairs were asked if they were satisfied with level their organization was engaged in program evaluation. Figure 14 below shows how many were satisfied and dissatisfied.

Figure 14 – Board Chair’s Satisfaction with Organization’s Level of Engagement in P.E.



Board chairs who were dissatisfied suggested what could be done to enhance the level their organization was engaged in program evaluation. Their suggestions included:

- Educate the board about evaluation (2)
- Conduct a consumer needs assessment
- Conduct a more in-depth evaluation with better evaluative questions
- Ensure program data is available more frequently
- Board should balance financial discussions with program related issues
- Staff should focus on evaluation as part of their work
- The ED should value program evaluation and make it a priority for the organization
- Develop meaningful outcomes to measure
- Formalize the evaluation process and ask the right questions

There are several reasons why half of the board chairs are dissatisfied with the level of their organization's engagement in program evaluation. The majority of these board chairs had previous instruction in evaluation and had higher educational levels indicating their dissatisfaction might be driven by a lack of rigor or comprehensiveness for the evaluations. Some of their comments that appear to support this included, "we only informally evaluate our programs;" "what we do I wouldn't call program evaluation;" and "we really only solicit informal feedback from our clients." The second most cited reason was that their organization didn't have the resources to adequately evaluate their programs.

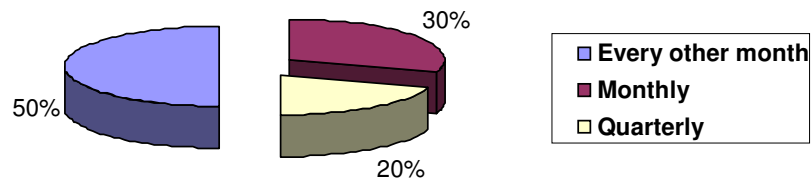
All board chairs were asked what the greatest challenges were for evaluating their organization's programs. Their responses included:

- Time – 8
- Cost – 6
- Expertise – 3
- Maintaining contact with consumers – 2
- Board's ignorance for program evaluation – 2
- Making sure we have meaningful outcomes
- Limited focus of board for program evaluation
- Having consistent data available
- Staffing
- Meeting multiple demands from funders
- Cultural change for long-time employees
- Projecting consumer needs
- Confidentiality (sensitive consumer information)

Figure 15 below shows how often the boards meet for the organizations of the 20 board chairs interviewed.

Figure 15 – Frequency of Board Meetings

FREQUENCY OF BOARD MEETINGS



Board chairs were asked how long they served on the board of their NHSO. They were also asked if they have ever received instruction in evaluation, not necessarily program evaluation, but any kind of evaluation. Lastly they were asked for their highest level of education. This information has been segregated out in Table 12 below between board chairs leading Type I and Type II organizations.

Table 12 – Board Chair Profile Information by Organization Type

RATING CATEGORY	MEAN *TENURE	MEDIAN *TENURE	RECEIVED INSTRUCTION** IN EVAL.	HIGHEST LEVEL OF EDUCATION
TYPE I	8.2 years	6 years	100%	20% - Doctorate 60% - Master's 20% - Bachelor's
TYPE II	8.2 years	6 years	53%	53% - Master's 40% - Bachelor's 7% - H.S. grad

n=20

*Tenure is total years served

**Instruction could be a college course, part of a college course, a workshop, a seminar, learning on the job or more than one of these examples

Table 12 indicates that the tenure of a board chair does not impact their organization's level of engagement in program evaluation, contrary to the executive directors. This is likely because of the differences in each other's roles within the

organization. Board chairs, while ultimately responsible for the organization, are volunteers and do not spend as much time with the organization's operations as an executive director would. It also indicates that the role of the ED as a driver for ensuring that programs are evaluated is much more influential than a board chair. The instruction in evaluation and education level variables are congruent with the results of the executive directors, indicating that previous exposure to evaluation in a learning context and advanced education is more likely to result in a board chair being familiar with program evaluation and possibly ensuring it is a topic for discussion with the board.

Program Staff

Program staff from 20 NHSOs were asked what program evaluation meant to them, and there was great variance among their responses. The majority however framed program evaluation within the context of delivering services to consumers. Within that context 40% of the respondents stated that it involved meeting goals or objectives, 20% referenced measuring outcomes, and 20% referenced measuring the effectiveness of the program. Other responses varied including "meeting program needs," "acquiring feedback," "determining how the program is delivered," "changing," "determining whether something is functioning properly," and "financial stability" among others. One fourth of the program staff talked about improving the program and/or services delivered to their consumers. A few talked about meeting the requirements of funders, and one described it as an ongoing process. Interestingly none at this point in the interview characterized it as a burden, something pulling their time away from consumers, or something they felt was unnecessary.

All of the counselors, psychologists, and therapists however described program evaluation as an administrative function and something they did not engage in or know much about. When probed further they talked about how they incorporated the latest research for their particular field; tested various methods and techniques; logged what was working and not working well with their consumer interventions; and tracked consumer progress, all arguably important aspects of program evaluation. Here is an example of such an exchange between the researcher and a psychotherapist:

- RESEARCHER: What would you say are the greatest challenges to having the program evaluated in this organization?
- THERAPIST: Oh God. I don't know. It's hard to answer that because I don't (pause) I'm not involved in that process.
- RESEARCHER: Let me ask you this. What would you say is the measure of success in this organization?
- THERAPIST: The measure of success (long pause).
- RESEARCHER: You know that what you're doing is working.
- THERAPIST: Successful treatment. We have, um, clients that they don't; they don't have to attend therapy here. We provide free therapy services.
- RESEARCHER: So, it's voluntary?
- THERAPIST: It's voluntary. Having that successful treatment, knowing that they've stayed and they've completed therapy. Regarding the forensic interviews, not every county uses child advocacy centers to do their forensic interviews, and so I see it just in the court system. And I hear it when I attend workshops and trainings with perhaps detectives that do forensic interviews, and I see their interviews. A lot of them are really not good because they can't rapport, and they can't get on that one-on-one with the child. So, I see it in that aspect.
- RESEARCHER: You said you know that it's successful if they have completed their therapy. Is there some sort of, for

lack of a better word, case management or record keeping, as a child would go through a certain period of therapy, that you would know when the therapy is completed, or successful at the conclusion, or –

THERAPIST: We have a case tracking system.

RESEARCHER: Okay.

THERAPIST: We document every session. We document all of the forensic interviews. You can pull up very quickly how long they've been here, how many, if we tried outreach, or even the number that we offered services to.

RESEARCHER: Okay.

THERAPIST: So we have a system of tracking that, but I'm not sure if that answered the question though.

RESEARCHER: Well, let's see. Let's say that I'm a child that comes here voluntarily for therapy. Can I assume that there is some kind of assessment up front?

THERAPIST: Yes.

RESEARCHER: And then based on that assessment there's some sort of recommended therapy that you would do based on that?

THERAPIST: Right.

RESEARCHER: Let's say it was nine months. At the end of the nine-month period, what would be an indicator for you that I'm leaving here now better than when I came before?

THERAPIST: We do a lot of therapy evaluations or assessments, or activities that we measure that we'll do in the beginning. Even just something as simple as a feeling chart, or feeling wheel to assess how they're doing. They'll read statements and check off how they apply. And we'll do the same ones after, maybe about two months before we think we're ending, you know, the end of therapy.

RESEARCHER: Okay.

THERAPIST: And, so we kind of, you know, go back. And we also base it on what parents are telling us - how they're doing at school, how they're doing at night.

RESEARCHER: Is that informal feedback, or do they fill out a form?

THERAPIST: No, it's just informal. Every time I meet with a therapy client, I also meet with a non-offending caregiver. And I do that on a weekly basis to just measure how that child is really doing

RESEARCHER: So, if I'm hearing this correctly, I'm new to this organization. I come here and you do some sort of assessment up front. I go through the therapy. You have checkpoints along the way to measure my progress, and at the end you do some sort of final assessment that you can compare to the first one. And through that, you also incorporate parent/parental feedback.

THERAPIST: Yes.

RESEARCHER: Okay.

THERAPIST: And we also keep that open. We recognize that the child was abused at age nine. When they are age thirteen, it could be very different because they have a better understanding of what happened to them. When they are nine they don't necessarily get that it was sexual, they just, they don't get it. But at thirteen, they're reaching that puberty, and sometimes it's a different ballgame for that. So we open, we keep those, you know, doors open that if you are having issues we explain to the child and the parent what to look for, uh, when they do need to receive therapy, if that's the case. And if they, if we have availability, we'll take that client back to kind of, uh, help them with whatever issues they're having.

RESEARCHER: You would have the file, of the history of what had happened?

THERAPIST: Yes.

RESEARCHER: Okay. A lot of the things that you are describing, the assessment up front, the tracking of progress along the way, the assessment at the end, and

receiving parental feedback, that's all part of program evaluation.

THERAPIST: Okay. Okay.

RESEARCHER: So, you are doing it.

THERAPIST: Okay. I guess when I hear it I just hear, you know, administration.

While the total number of staff from these professions in this study was small, this type of misunderstanding or mischaracterization should be considered for further study.

Program staff characterizations of their roles in program evaluation

The 20 program staff varied across hierarchical levels within their organizations, job titles, levels of responsibility, and specific tasks. Some examples of their job titles, that also reflected the diversity in service delivery among their NHSOs, included:

- Director of Campus Life
- Clinical Director
- Director, Workforce Services
- Case Manager
- Vice President, Program Services & Human Resources
- Employment Services Representative
- Programs Manager
- Association Director, Youth & Teen Development
- Program Manager
- Advocacy Coordinator
- Therapist, Program Coordinator

This variation apparently impacted their roles in the program evaluation process. Some are intricately involved in the process, some are not involved at all, and some are program directors who are responsible for making sure the evaluation of the program takes place. Some examples of responses illustrating this variation include:

- “Being a manager, I would be the one to either observe the situation or the problems in case it looks like maybe there needs to be a change here and there. I would relay that to our management team where we would discuss it and many times it would be a couple of the managers

involved and of course the frontline people who deal with this everyday.”

- “Well, I mean it’s our job as directors of the department, you know, to be sure that care managers work with their clients and meet their needs, follow up on concerns, and do all of those things as we direct the department”
- “It’s usually a shared responsibility between the executive director and myself, the program manager.”
- “I am the Director of the Institutional Programs. There are three different service programs attached to that. So I am the person that is the contact person for those programs to provide the data that is required.”
- “My job is to work as a consultant to all of the program managers. If they have evaluation dollars within their budget, they’ll come to me and we’ll craft through what an evaluation would look like and how to write the RFP. Sometimes I sit in and, for bigger evaluations, will be part of the evaluation selection committee.”

The delineations in the tasks they performed in their roles were not clear cut. The majority of them performed a mix of tasks and responsibilities related to program evaluation. For example, one was responsible for the peer review process for her members while someone else was responsible for developing and administering a survey to those members. Another stated, “I develop the questions, the overall format, and then consult with our staff to find out if there are any concerns that they would like brought into the evaluation process.”

The majority of program staff emphasized the importance of acquiring consumer feedback and using it to change aspects of their program. Some incorporated feedback loops with their consumers and expressed the importance of their role in the evaluation process. Examples of some of their responses included:

- “The activity program and the activity evaluation, like I said, it’s not about me. It’s about what the residents want to do. And, we try to accommodate that by whatever means I can.”
- “I ask for feedback from clients in a way of holding them accountable and wanting to know how they think they’re doing. I also believe in

having them participate, setting their goals, and reviewing themselves just kind of, again, part of therapy mode.”

- “I will be the one that will go back to the kids and communicate to them the results of the survey, what things we’re looking at changing, and anything on there that we just know we can’t change.”

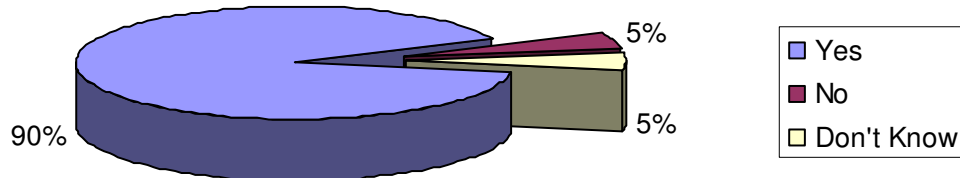
Some indicated they were not part of the evaluation process by describing their role as delivering the program, indicating indirectly that the evaluation function was either limited or performed by someone else. Only one described how she and her coworker, the other program director, were resources for the evaluators their organization hired. This appeared to confirm from the 42 ED interviews that the majority of program evaluations in these NHSOs were conducted internally. She explained, “I’d say our primary role in this evaluation is to let these evaluators know what our main responsibilities are so that they’re honing in on evaluating that piece of it and then the result of that.”

Program staff characterizations of their executive director’s roles in program evaluation

Program staff were also asked whether their EDs had a role in their organization’s program evaluation process, and if so to characterize that role. Figure 16 below shows the percentages of responses from program staff.

Figure 16 – EDs Having a Role in the Program Evaluation Process

**EDs HAVING A ROLE IN THE PROGRAM EVALUATION
PROCESS - according to program staff
n=20**



There were no clear differences between responses from program staff working at Type I organizations and those that worked at Type II organizations. Here is the breakout of how 90% described their roles along with some examples of their responses.

Overseer – 10
Driver – 4
Facilitator
Advisor
Collaborator
Supporter

- “I believe that was to ensure that the members were encouraged to participate fully, that they not feel threatened by their participation, and that the information they shared was not going to be used against them.”
- “We pretty much share the same responsibilities. She also meets with judges and judicial staff for more in-depth information after any concerns are expressed.”
- “He doesn’t have an active role in the process but yes he does have a role in paying attention that we’re doing it and what are the outcomes. You know he wants the information, he wants to know what’s going on. These things are reported at our monthly LT meeting. All the outcomes are on the Board report, not all the outcomes but some of the main ones are on the Board report that

goes out every month and so he sees all that and kind of knows what is going on.”

- “He has access to all of the reports and I don’t report directly to the Executive Director, the vice president does. Now I will tell you on the older worker program because it is a national grantee and XXX International is involved, he is getting those reports and I’m sitting down with him quarterly going through that. So I think the programs that have more visibility to him, he’s involved in. The programs that have less visibility, he’s not involved in.”
- “Well she’s the one who looks at it and then I guess accumulates the information and I think gives it to the Board. That’s what I think happens but I’m not sure.”
- “Yes. He is one of probably five or six of us who serve on the Quality Assurance Committee. He directs the proceedings of the QA Committee and he has specific responsibilities that have to do with measurement of risk factors and insurance related factors. I think he is overall the person that I guess makes sure the rest of us do what we need to do in a timely fashion.”
- “He does, and he may do more than I think he does with it, but he always reviews the outcomes and highs and lows and wants to know about corrective actions or improvements. ‘What are you doing right? Let’s see how we can keep that and what are we doing wrong? How can we fix it?’ And he’s been a driver in some of the survey formations, an outside group doing it, but from my chair generally we share the results with him and he’ll sit down and talk through.”
- “He’s apprised of the results that are presented to our staff and our board of directors. You know, so I mean does he have direct involvement? Maybe not, but he’s certainly aware of results and trends.”
- “Demanding them (laughter). She wants evaluation results, and I’m thinking of particular program evaluations. She certainly dictates, lays forth a vision for the organization, and has led this organization in the process to develop our own internal metrics for success, like a scorecard type. So certainly she is very tied to the outcomes of the organization.”
- “I think she leads it. I mean I think she designs anything that occurs formally at least.”

Program staff participants were given the opportunity to talk about any challenges program evaluation presents for their jobs. First, they were asked if the evaluation process presented any trade-off (opportunity) costs with their time,

as in lieu of what they would normally be doing such as serving their consumers.

The breakout of program staff indicating whether there was a trade-off cost for their time is broken out by those working for Type I and Type II organizations below in Figure 17.

Figure 17 – Trade-Off Costs for Program Staff by Organization Type

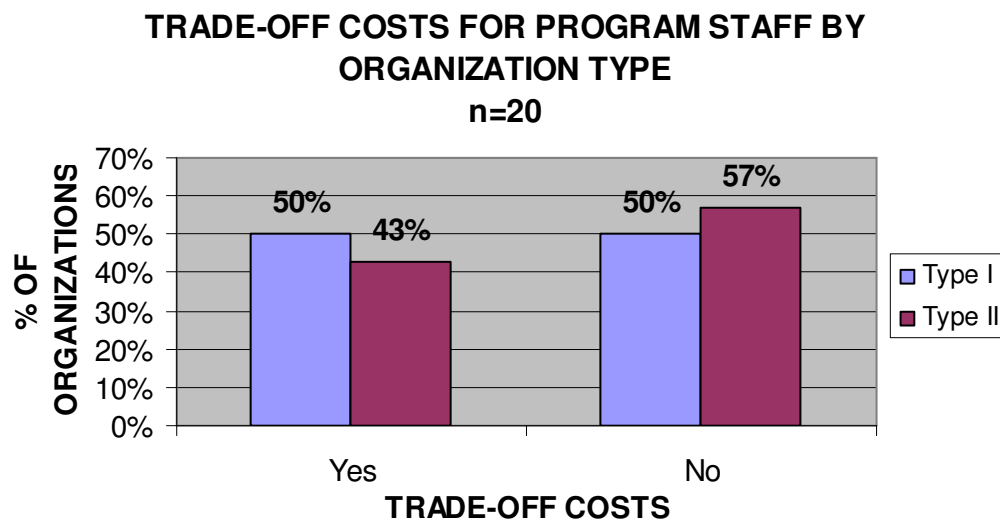


Figure 17 does not tell us the entire story for trade-off costs because it is important to understand the context and environment behind their responses. The program staff for Type I organizations conveyed various reasons for why there was a trade-off cost for their time. Two program directors in a large organization stated that the trade-off would be their time which could be spent supporting and training their care managers. However, they also indicated the process was worth the time, and presented the researcher with a 30-page outcomes report that they seemed proud of and discussed at length. The Vice President of Program and Youth Development Services for a national headquarters of a NHSO indicated that her program directors would “certainly be doing

other things,” however also stated that “I think we’ve come to the point where it’s just sort of a built in expectation.”

The program staff from Type II organizations that indicated they had trade-off costs for various reasons. One conveyed that it was a staffing capacity issue and “There are some things that have to be set aside.” Another indicated it simply took time away from program staff normally spent serving additional consumers but that the organizations had been successfully streamlining the evaluation process to minimize the trade-off. Others characterized it as a burden driven by external and internal reporting requirements as reflected in their responses:

- “I would say without a doubt in the last couple of weeks, because we’re approaching the end of the first quarter. It’s record review time so these people are going to review those charts and those people over there are going to review these charts, so that you have an unbiased look at your client files in order to make sure that you’re collecting data you’re suppose to. You are looking at a deadline of getting all the records reviewed, collecting the data, putting it in some fashion for the quality assurance person, then developing an action plan based on a comparison with the last quarter. And clearly the hours put into that or days put into that would have been time spent working directly with consumers or providing supervision to staff that do work directly with consumers.”
- “For the most part not. But, I have to say the parts that have to do with reporting to the State, the amendments to the grant, and all the paperwork that it takes involved do. A lot of that is outcomes reporting. It keeps you from really putting some of the other things about programming in place you’d like to. The Director of the Itinerary of Rehabilitation Services and I spend a lot of our time writing these things or sending in the reporting rather than really looking at what is going on, where we really need to be, and what do we need to do to move there.”

A representative of a youth-serving organization stated it also presents a trade-off cost for the finite time available to communicate with volunteers, such as when a portion of volunteer meetings is spent on discussing the evaluation instead of other topics. But, he

concluded that, “It’s probably going to make the next program more successful so you know, it’s a valuable tradeoff.”

The program staff from Type I organizations stated there were not trade-off costs on their time because program evaluation was the responsibility of management or it was an integrated part of their work. For example, one representative of a youth-serving organization stated, “It’s part of what their expectations are. You know we don’t do it every week but it’s something when survey time rolls around, I know that that’s what my responsibility is for it.” The other staff that conveyed they didn’t have trade-off costs because it was part of their job represented, upon review, what might be considered “low level” Type II organizations. Their evaluation efforts are a bare minimum and typically include using only a customer satisfaction survey, informal consumer feedback, or their own observations of what transpires to report problems or successes. A representative of an organization that serves senior citizens looked at her time spent in the evaluation as an investment for saving time spent on future evaluations.

In a different context, a psychotherapist said, “Right now I am in the position where seeing clients is requiring all of the twenty hours that I am funded to work here. So, I would not trade off seeing clients for program evaluation. That’s probably unfortunate, but I couldn’t justify that.” When she was probed further in the interview she mentioned parts of her work that are elements of program evaluation although at first she didn’t think they were. They included tracking consumer progress, determining when the cases should be terminated, revamping her approaches and methods based on her case notes, and incorporating the latest research-based recommended practices.

Program staff participants conveyed what their greatest challenges were for the evaluation of their programs. The majority of participants cited multiple challenges, some specific to their environment, organization, or their work. These challenges fell into various categories as shown below.

Knowledge, skills or expertise

- Understanding how to evaluate the programs of a membership agency
- Having qualified staff to conduct phone surveys and accurately glean information from clients
- Staff not understanding that the residents should be holistically approached
- I would say the biggest challenge is that for a lot of things that we're doing for the first time we simply don't know if we're overdoing it, if we're maybe a little bit off track in why we're doing it.
- Knowledge for a comprehensive evaluation process

Stakeholder involvement

- Not being involved in all program areas
- Getting volunteers to participate in the evaluation process
- Probably the stakeholder input, because, we struggle with that
- Oh, God. I don't know. It's hard to answer that because I don't, I'm not involved in that process.
- Not receiving the evaluations of my service to our consumers.

Resources

- Time – 7
- Funds – 2
 - For the evaluation process
 - To hire professional program staff

Structures

- We're trying to develop a database that is driven by a web based data entry program to support our evaluation efforts.
- Well frankly, right now it's ownership. You know, is it going to be my responsibility to do these? I just happen to do it because I've been around a while and ask if we're rolling out the after school survey evaluation. Then I get asked to do it.

Process

- Communicating evaluation results to who needs to know our challenges and what we do well

- Adjusting to constant program change due to changing government regulations (USDA)
- The multiple aspects of evaluating our services and the volume of our consumers present the biggest challenges. I have to look at where they are working, what they are doing, how much they are getting paid, and determine whether it is better for them to be working or sitting at home collecting their social security.
- Streamlining the evaluation process to do it effectively and efficiently
- Feeling that it's a choice between providing services, and assessing them
- Trying to really make sure that we're identifying the things that we should be measuring and we're not just coming up with measures to say we have measures. You know I want to know that what we're measuring makes sense to the quality of the program and that it really is going to keep our programs.
- Data collection at the local level. Identifying the metrics that are important and are measurable, and then collecting that data.
- Accepting the limitations of any evaluation process. We want to make sure they're in our control, that we influence them, that we can measure them, and sometimes when you put all of your framing on it you get left with participation numbers. And you don't want to, because sometimes that doesn't feel good enough and in some cases it's not good enough.
- Consumer fatigue for participating in our evaluations. They get tired of being asked these questions. That's definitely an issue.
- We have a transient population so for those in inner city schools the surveys come back in return mail. So we try to find other avenues of how maybe we do a family night at that school and we ask while they're there, ask them to fill it out that way.
- Communication between the staff and our board. You have some staff who are in the board meetings, but there's never really straight communication. There's not too much there that the staff get to do or that the board gets to do with the staff. So they can be seeing totally opposite things.
- Getting our consumers to complete our surveys

Culture

- The lack of staff's ability to be candid about it
- Getting staff on the same page
- Having some of the old school staff (especially old school nurses) go beyond the physical aspects of care to integrate those that are spiritual.
- Acquiring honest responses from a balanced, representative sample of our consumers. We tend to hear the ones that complain the loudest or the ones that appear to always be pleased because they don't want to offend anyone.

- The greatest challenge to that is because as the case manager I want to see everybody succeed, ready to always to work. It doesn't always work that way.
- Getting our ED to make it a priority
- When staff take the evaluation results personally. It's generally minimal but you'll hear some frustration from folks. Primarily it's because volunteer or paid people put their heart and soul into it and perceived negative comments in an evaluation hurt them.

Internal demands

- I would say probably just meeting the milestones, you know meeting that quota each month. I just want to meet those milestones and if so, go over and beyond not for personal reasons but because if I'm not meeting them then that is how many more consumers that are not going to be finding jobs. That means I'm not doing the service that they came for me to do.
- Satisfying our consumers by making sure I can use all of the input and find a reasonable solution to fix it that fits everywhere. Because you can always change a program one way and the other side's going to hate it. It's a balancing act.

External demands

- Maybe if we had already become accredited and there's certainly a matter of maintaining rather than getting things done for the first time, maybe there wouldn't be so many different things.
- I think the greatest challenge is that the programs don't all run the same time of year. It depends on the funder and keeping up right now with who should be called and when. We have a senior program that is funded from October 1 to the end of September every year but the rest of our programs are pretty much from July 1 to June 30th and then we have the summer program which is just the four weeks and trying to keep track of when somebody finished a program and then again in a year is just ridiculous.
- The other thing is the funding aspect. Right now I don't think we do a good job of figuring out how much it costs to serve in some of the programs. I mean yeah we get funding from the State or we get funding from this but because of our database we are not able to say this person was served from here to here and this is how much it would have cost.
- I think we've struggled with having a diverse funding base. Some of our funding streams require one level or one element of evaluation differently than others. CARF wants you to do it this way and that's how we are going to do it in this part of the organization, so trying to come up with that consistent overarching program evaluation strategy regardless of the funding stream. And, that goes beyond program evaluation. It goes to program design, you know, we're trying to

move towards we're going to determine what our service delivery model is and then finding sources to support that model rather than finding the funding source and then developing the service delivery model.

These challenges were not broken out by organization type due the great variance across all 20 participants and Type I and II organizations yielding no clear trends.

In spite of all these challenges, all 20 program staff stated they would still engage in having their programs evaluated if they were given a choice. Some examples of their responses include:

- “I would still do it because I would still want to know how the kids feel, they're treated and what they like and don't like. And I would want the same information back from the staff.”
- “Well because I'm really invested in doing good work and providing quality services.”
- “I do think that evaluation adds a level of discipline to a program implementation. On the front end you have to state what you're looking to get out of it. You run the intervention of a program, and then you have to collect data either along the way or at the end to see if you met your aims. And the reason why we do any of these programs is because we think it's good for kids, and that they are impacted by them.”
- “Frankly I would pick continuing to do it but finding a cheaper way to do it”.
- “I think it makes the parents realize that we're being held accountable and that they have some input into that accountability.”

A common theme for the majority of the participants when this was discussed was that they also recognized that program evaluation was important because their environment and their consumers' needs are not static, therefore their programs also must change.

Program staff were asked how long they worked at their NHSO. They were also asked if they have ever received instruction in evaluation, not necessarily program evaluation, but any kind of evaluation. Lastly they were asked for their highest level of

education. This information has been segregated out in Table 13 below between program staff working at Type I and Type II organizations.

Table 13 – Program Staff Profile Information by Organization Type

RATING CATEGORY	MEAN *TENURE	MEDIAN *TENURE	RECEIVED INSTRUCTION** IN EVAL.	HIGHEST LEVEL OF EDUCATION
TYPE I	10.3 years	9 years	83%	14% - Doctorate 43% - Master's 43% - Bachelor's
TYPE II	8.2 years	4.5 years	71%	7% - Doctorate 14% - Master's 50% - Bachelor's 29% - H.S. grad

n=20

*Tenure is total years served at the organization

**Instruction could be a college course, part of a college course, a workshop, a seminar, learning on the job or more than one of these examples

The program staff's tenure, instruction in evaluation and level of education all appear to have a relationship with the category type of their organization. This information is consistent with the results for the EDs and the instruction in evaluation and level of education variables in board chairs.

Funders

A total of 24 funders that fund NHSOs were interviewed comprising 4 (2 from each MSA) each from the United Way, community foundations, corporate foundations, independent foundations, corporate giving programs and government agencies.

Here is the summary breakout of how these funders described what program evaluation meant to them:

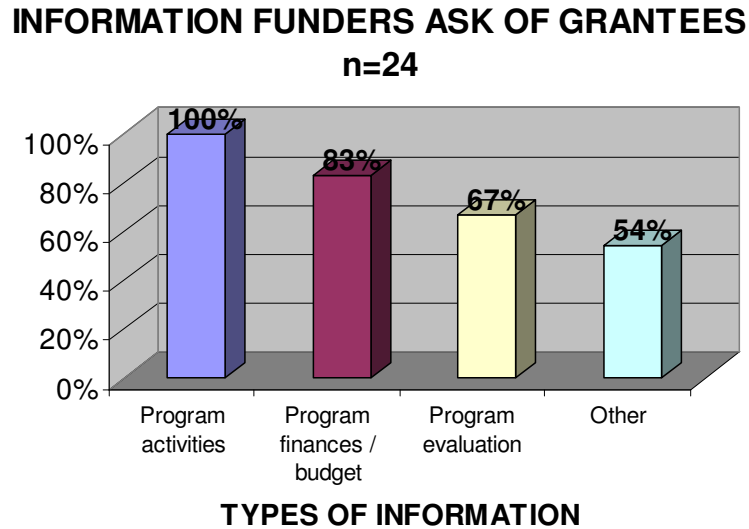
- Evaluation of the impact of the grant – 4
- Evaluation of the grantee organization – 2
- Looking at the impact of the program on the people it serves – 2
- How effective the services are that they are providing to the community and to the clients that they serve – 2
- Determining if a program has met its objectives – 2
- Different things to different people – 2

- Could include various processes including auditing, compliance, oversight, building knowledge, learning, improvement, social change or a mix of these.
- Depends on who's talking about it – are they referencing a program activity versus, impact analysis as it relates to, in effect, the outcomes and especially the long-term outcomes of a particular program initiative, etc.
- Assessing where the programs are that we have supported
- A grantee organization's internal testing process
- Examining how the program works
- A systematic, integrated process of planning, designing, implementing a program and measuring the program's outcomes
- Evaluating a particular component or program that the grantee operates or offers
- Determining the success of the program
- The quantity and quality of services including indicators, benchmarks.
- Process and outcomes. Determining if the process has been efficient and effective. Determining what the outcomes of the program are compared to what is intended in the program's logic model.
- Measuring indicators to ensure that appropriate services are being provided

We see here a wide variety of descriptions, some of which appear related to the context, what is considered the “program,” and the perspectives of the people attempting to define program evaluation. Two trends emerged from these 24 responses. First, the representatives of corporate foundations and corporate giving programs gravitated to the evaluation of the impact of their grant or of the organization they are funding. Second, the descriptions most closest to textbook definitions come from people who are professional evaluators or were in previous employment capacities.

Several questions attempted to find out what these funders ask of their grantees. Figure 18 below shows the percentage of funders that ask their grantees for several types of program-related information.

Figure 18 – Information Funders ask of Grantees



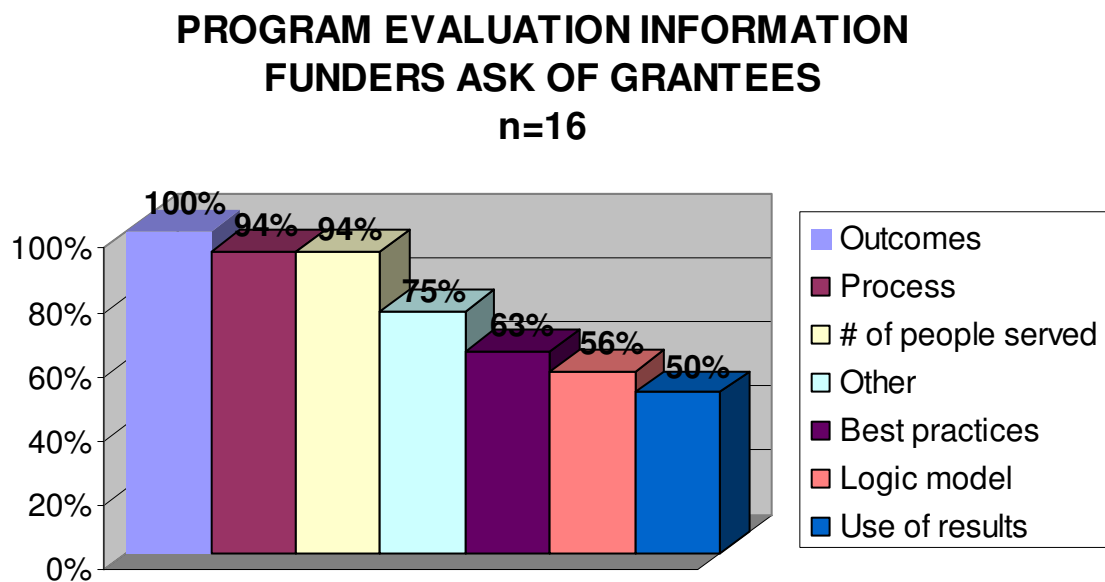
The responses for those funders indicating they asked for “other” information beyond the three choices given varied greatly:

Other

- How the money was spent – 2
- Company’s employees pick the programs our foundation funds usually based on where they donate, volunteer, or serve on boards, and they provide information about those programs to the company
- The end results of how our funds were used
- Information on the stability of the board – fiduciary responsibility, ability to govern itself, etc.
- Information to ensure programs are in compliance with federal guidelines
- Who has already supported the organization and how much
- Detailed operating budget with audit statements
- Where program and recipients reside
- How the organization has or plans to leverage other dollars
- Explanation of how the evaluation fits the work of the program
- Specific questions about the programs
- Board roster and board contributions
- The program’s specific goals or objectives and how they fit into the larger strategic direction of the organization
- Specific outcome data

The 67% of funders that stated they ask for program evaluation information were asked to indicate within that category what types of information they ask of their grantees. Figure 19 below shows the percentage of funders that ask for six types of program evaluation information along with their responses for other types of information.

Figure 19 – Program Evaluation Information Funders ask of Grantees



KEY: Outcomes	=	program outcomes
Process	=	program process (how it is delivered)
# of people served	=	# of program's consumers served
Other	=	other program evaluation related information
Best practices	=	service delivery or profession's best practices
Logic model	=	program's logic model (program theory)
Use of results	=	program evaluation results are used

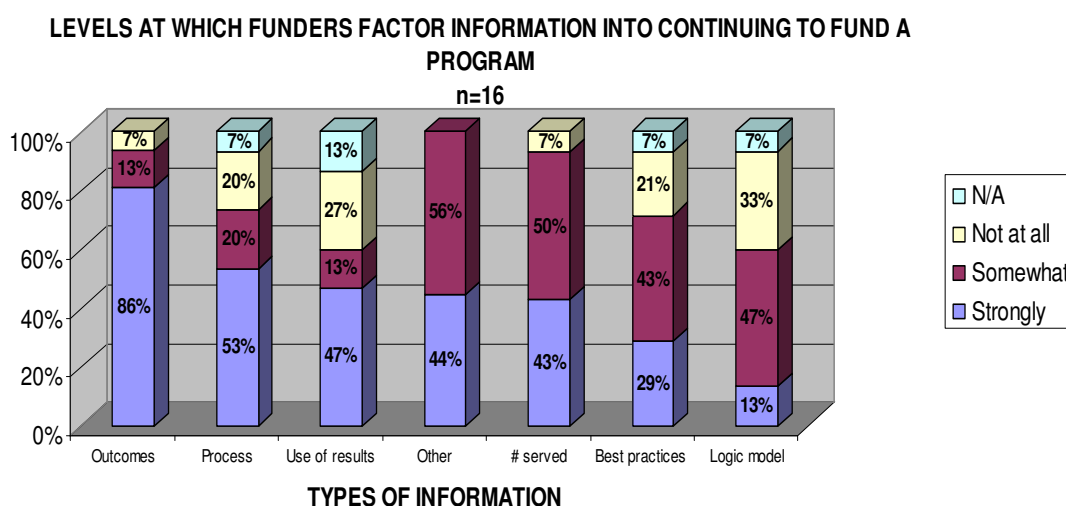
Other

- How the program's impact fits with our strategic initiative – 4
- How the program benefited the community – 3
- What worked and what didn't – lessons learned – 3
- Other funders or accrediting bodies requiring evaluation information
- Consumer satisfaction
- Defining success

- Outcome targets
- Demographic area where program was evaluated
- How does the board assess what programs should be added or deleted
- What is the board's interest in the program
- Evidence that program's process was compliant with our regulations

Funders were also asked how significant the program evaluation information they ask for factors into their decision to continue funding the program. Figure 20 below breaks out each type of information by their level of importance in this decision.

Figure 20 – Levels at which Funders Factor Information into Funding a Program



A few funders indicated there other types of information that factor in their decisions whether to continue to fund programs. These responses are listed below.

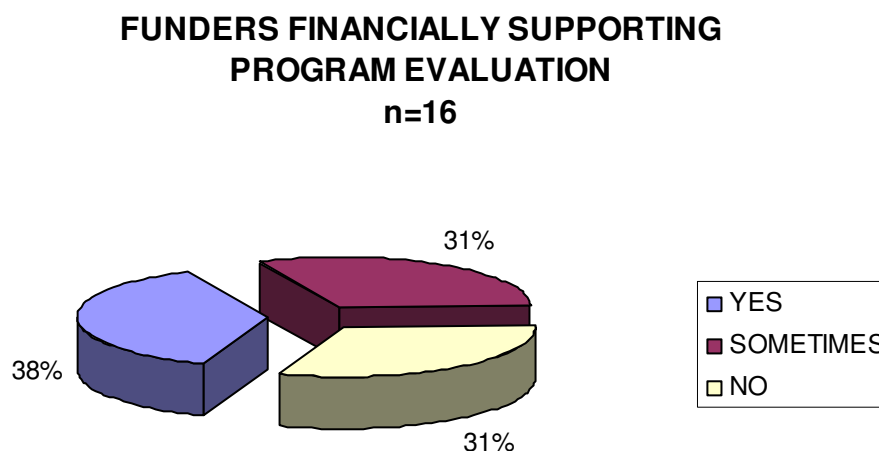
Other

- Community needs assessment – strongly
- Employees requests for funding programs – strongly
- Whether the organization was compliant with our regulations for the program's process and performance measurement – strongly
- How the organization's program measured up against our points scale – strongly
- Lessons learned, positive or negative – somewhat
- Their ability to use our grant to leverage other resources – somewhat
- Innovation – somewhat

- Leadership – somewhat
- Strength in working in collaboration – somewhat

An important factor in ECB for NHSOs can be the financial support from their funders, particularly those who require programs to be evaluated at some level. The 16 funders requiring program evaluation information were asked if they included financial support for the evaluation of funded programs, and the results are shown below in Figure 21.

Figure 21 – Funders Financially Supporting Program Evaluation



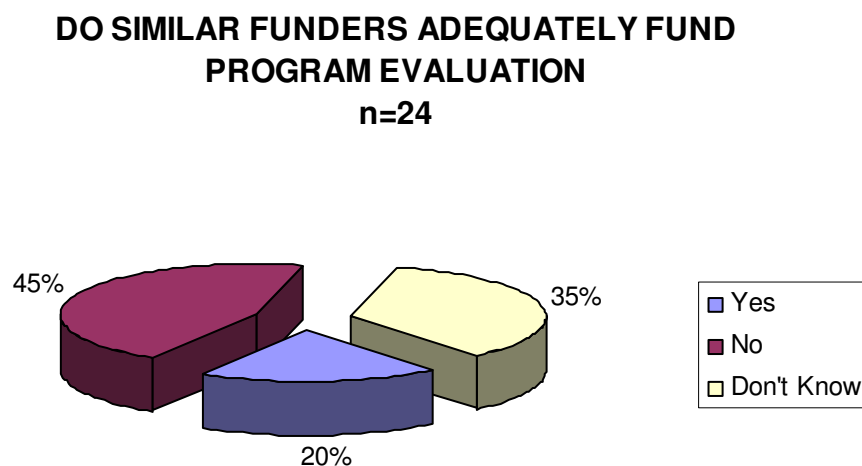
Some of the funders who stated they sometimes funded program evaluation gave examples for appropriate situations:

- If they are doing something that we would consider new and innovative that we would like to understand better and see if it's worth encouraging others to replicate. And I wish we were doing more than that.
- It depends on what initiative they fit in and how much money we have in our capacity building fund. It's not available to everyone. So usually it's those who are most in trouble that we might entertain that kind of a request from.

The funders that do not fund program evaluation were asked whether they would consider it if a grantee asked for the funds or included them in a grant proposal. The

majority of them (80%) stated they would consider it, but they also indicated that no one ever asks for it. The funders that do fund program evaluation determine how much to fund it in different ways. They include having a set amount for evaluation regardless of the program or grant amount; a percentage of the grant amount typically 5% or 10%; determining the amount based on what the program requires and having the evaluation bid out for contract; and others. All 24 funders were asked if similar funders i.e. other community foundations for a community foundation adequately fund program evaluation, and Figure 22 below shows how they answered.

Figure 22 – Opinion for Similar Funders’ Support of Program Evaluation



Debate has existed on the effects of foundation support regarding the possible enabling of dependency on foundations with long term funding versus the inability of programs to be sustainable with short term funding. The Rooney and Frederick study of foundations’ funding of overhead costs showed that foundations were willing to fund overhead costs within a grant proposal. The study also revealed that most nonprofit organizations do not rely on foundation funding to pay for their core operations (2007). These two results have implications for the funding of program evaluation, but some

clarity and consistency among all funders and NHSOs as to whether such costs are program costs considered “core operations” or “overhead costs.” The funders in this study who do not fund program evaluation but who also stated they would consider it if asked, but were never asked, indicates that the grantee organizations may be contributing to the lack of financial support. The majority of those funders who are currently funding program evaluation costs indicated they wished they could do more.

Table 14 – Funder Profile Information

*MEAN TOTAL ANNUAL GIVING	*MEDIAN TOTAL ANNUAL GIVING	RECEIVED INSTRUCTION** IN EVAL.	HIGHEST LEVEL OF EDUCATION
\$28,124,713	\$2,300,000	91%	14% - Doctorate 41% - Master's 41% - Bachelor's 5% - H.S. grad

n=24

*Amounts are estimated – fiscal year 2006-2007

**Instruction could be a college course, part of a college course, a workshop, a seminar, learning on the job or more than one of these examples

The variance in the total annual giving is due to the diversity of the 24 funders that participated in this study. Organizations varied in size and type such as community foundation, United Way, etc. The independent foundations ranged from some of the largest in the US to some of the smallest, while corporate giving programs varied due to their giving amounts being pre determined by the company’s CEO, a percentage of previous annual earnings, or driven by the employees. The variance in education levels is most likely driven by the variance in size, as in large foundations program officers are responsible for a larger number of grantees, projects, initiatives and allocated dollars to be granted. Smaller, independent foundations, which mostly are family foundations, typically have one person administering the grant process with amounts determined by the board which usually is the family. Quite often, family foundations contract with financial managers who have multiple client portfolios they are responsible for.

Evaluators

A total of 20 evaluators that evaluate nonprofit human service programs were interviewed comprising 10 each from the Atlanta and Indianapolis MSAs. This group represents the diversity in evaluator work environments including independent consultants, internal evaluators at large NHSOs, consulting firms or management support organizations (MSOs), government, and academia.

Evaluators' descriptions for what program evaluation meant to them overall were comprehensive and most took a holistic approach to the evaluation of a program. Most of their descriptions also reflected the intensity they displayed throughout the interviews and for their profession. As expected, their descriptions on average were six times longer than the other four stakeholder groups. Theirs would be in many cases entire or several paragraphs, as in the two examples below:

- “I think it has to do with looking at the effectiveness of the program. I think it can also have to do with the -- depending on the stage at which evaluation takes place, it could do with the process of determining what the structure of the program is going to be, what its goals and objectives are going to be, and then the implementation of what is proposed to be theoretically implemented, and then a documentation and the observation of what actually ends up being put in place, which we know often differs.”
- “I work currently as the primary evaluator on three programs, and they all are in the social services. Two deal with with substance abuse prevention and one is homeless intervention. And so to me that means monitoring programs, evaluating the outcomes on a monthly, quarterly basis, providing feedback to not only the grantor but also to the grantees, to the participating parties, interacting with staff, and doing a complete analysis of not just the hard data, but also evaluating the process that they use. So you always have the two components, the outcome evaluation and the process evaluation. I try to get most people think of the outcome evaluation first, that kind of intuitiveness where people know you collect data and they receive reports from you. But I also try to stress and emphasize that the process evaluation is almost if not more important in some programs. So we always try to provide a little bit of a quality feedback loop to the programs we work

on, since they are for the most part running between three and five years. So that gives you a little bit of time to tweak things. Sometimes you have to tweak a lot. Others it's just a little correction here and there. But in the end it helps the program provide better results, better outcomes. And then this day and age where everybody gets evaluated and everybody -- their performance is measured, and it becomes a part of your record, especially with federal and state funding, that's an important component we can provide."

The emphasis and focus on the program was a central theme throughout the interviews as the word "program" appeared a total of 65 times in the 20 transcripts. As one evaluator put it, "And there's lots of confusion out there in the field, the field even recognizes it. The evaluation field recognizes this probably more so than the field of assessment. So program evaluation is evaluating the program, underscore, emphasis, *italics* - program." "Evaluation" was the second most common word referenced 32 times reflecting their expected comfort with the word that embodies their profession. These interviews represented a desire to explain program evaluation as a systematic process, drawing from Patton's definition as in the examples below:

- "Program evaluation. I often think first about a definition that we use often. It is the systematic collection and analysis of data in order to make decisions. So I think about systematic data collection in order to make decisions. That there are decisions to be made, questions to be answered, and that it's often related, if it's program evaluation, related to a program, project, something along those lines."
- "I think of Michael Quinn Patton's definition when I think of program evaluation."
- "You'd think I'd have this memorized. I've said this at so many classes. Systematic investigation of a program's merit or worth, I guess. I'm probably combining Patton and Scriven there."

The term "process" was used to describe program evaluation in 14 of the 20 interviews.

A review of four of the other six transcripts shows that while they didn't use the term,

their description for program evaluation described what we would consider a process, rather than a task, product or report.

The majority of their descriptions had Patton's utilization focus incorporating phrases such as:

- "Giving the organization information they can use to improve themselves"
- "Trying to help people get information to make good decisions"
- "Determining what the structure of the program is going to be, what its goals and objectives are going to be, and then the implementation of what is proposed"
- "For me it's a whole process of improving, insuring we have programs and projects that are of the highest quality, the intensity and duration that we need to have, the kind of program impacts we want to have, and outcomes we want to achieve."
- "Sometimes you have to tweak a lot. Others it's just a little correction here and there."

A discussion about ECB comprising several questions and follow up probes took place in every interview with the evaluators. They were asked to define evaluation capacity, discuss the evaluator's role in ECB and share some of the challenges, successes and techniques used for effective ECB.

A categorical summary of the main themes that were part of their definitions for evaluation capacity is shown below:

- ECB is a continuum
 - Evaluation plan
 - External environment
 - Support
- Knowledge and awareness of evaluation
 - Data collection – 4
 - Methods – 3
 - Implementation – 2
 - Coaching stakeholders – 2
 - Training stakeholders
 - Evaluation language
 - Program theory
 - Logic models

- Data analysis
 - Evaluator's knowledge, skills and experience
- Resources
 - Personnel – 8
 - Funding – 7
 - Time – 6
 - Technology – 2
- Organizational culture
 - Commitment – 12
 - Staff
 - Board
 - Leadership – 4
 - Prioritization – 3
 - Stakeholder participation – 3
 - Enthusiasm – 3
 - Cooperation – 2
 - Understanding importance
 - Alignment
 - Organizational context
 - Moving beyond external demands to intrinsic motivation
- Organizational learning
 - Utilization – 4
 - What's working
 - What's not working
- Institutionalization
 - Evaluation is a basic assumption
 - Evaluation is a priority at the same level as other business functions

Some examples of their responses for their definitions include:

- “I think the ability of an organization to conduct its own evaluations and, to a minor extent, the ability of an organization to understand, take action on evaluation findings produced by others.”
- “Has the organization internalized the evaluation? Is it part of every day business? When evaluation is on the same priority level as all those other things – financial management, human resources, etc.”
- “One would be the expertise, somebody who know how to do things and what to do, to lead people through the process. There's I guess financial resource capacity and that can include funding for staff time, funding for resources and really growing as you need to.”
- “The commitment to the process is necessary. Even if you've got the means and all of the rest, if your staff and your board aren't committed to the process, it's not particularly useful to anybody.”

Evaluators were asked if they had a role in ECB with NHSOs and if so, to describe that role. All 20 evaluators said they have a role in ECB, with slightly more than half indicating they feel it is an inherent part of their job and an obligation. Some even go so far as to state that ECB involves an external evaluator effectively working themselves out of a job, so the organization can conduct their own evaluations. They characterized an evaluator as someone who wears many hats and can play different roles depending on the organization's stakeholders, culture, needs and wants for program evaluation. At any given time an evaluator can be a collaborator, convener, facilitator, trainer, presenter, mentor, coach, salesperson, amateur psychologist or sociologist, project manager, or educator.

There are some steps evaluators can take in the ECB process that were commonly referenced by the evaluators that were interviewed. For example, they can form collaborative initiatives made up of similar organizations and facilitate technical assistance in a setting optimum for peer learning. They can bring their expertise and impart wisdom to staff. They can get stakeholders to think about evaluation when they develop new programs. Helping stakeholders develop their theory of intervention for their programs was consistently cited as an important step evaluators can take for ECB. The evaluator can be the person who can assess an organization's knowledge and technical capacity for program evaluation at the beginning of an assignment. This involves looking at the organizations strengths such as what data they are collecting, what they could be collecting at little or no extra cost, what infrastructure needs the organization has to adequately engage in program evaluation, and even help allocate resources for the evaluation.

Evaluators are often seen as outsiders by an organization's stakeholders, but they can have a role in developing an organization's culture for evaluation. An overarching theme in all of the interviews was the importance of evaluators helping stakeholders understand what it is they're trying to do, how they're trying to do it, how to organize it and to track it. Evaluators can also strive to empower stakeholders to realize their potential for their part in the evaluation process as well as help them think evaluatively to realize opportunities for change and improvement. Three examples below capture several ways an evaluator can play an important role in ECB.

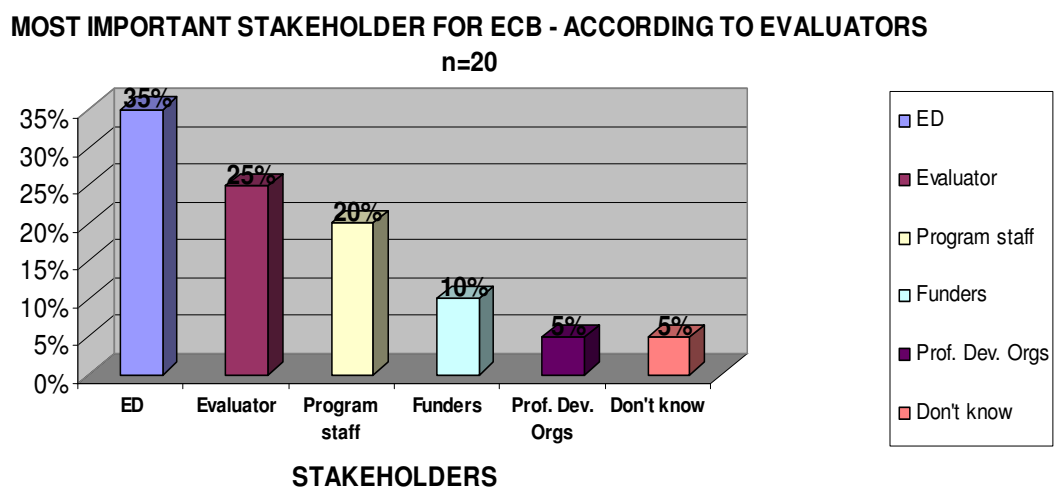
- “I played that role in a multi service center that had many different kinds of programs including for the elderly, youth, emergency services, people in need, people in the substance abuse program, programs for runaways, and things like that. I developed a very simple process by which as staff members, we drew up kind of a summary of each of our programs. They included their mission, funding, service goals, and staffing. Then we assigned our board members to go into those programs and do a very cursory kind of program review than a program evaluation. But probably more than that, it was a great way of educating your board.”
- “One of the things we developed is an evaluation capacity assessment tool that we use. It asks them about a variety of things including access to resources and things about commitment. You know, do they have access to persons to help understand statistics or analyses? Do they have people in place to utilize, information somehow within the organization? Do they have basic computer skills or computer tools?”
- “It seems to me as an evaluator, when I go in there I would want to as much as possible to try and engage the clients to a point where I'm leaving a legacy they can pick up on. So, you know, try not to create one-time surveys but surveys that can be legacy instruments. Try not to create one-time, sporadic data collection sources.”

The evaluators were asked if there was anyone else within the context of evaluating nonprofit human service programs that could have a role in ECB.

Overwhelmingly, they stressed the importance of stakeholder involvement and the fact that they could not successfully build evaluation capacity by themselves. They discussed

who some of the other players might be and some specific steps they could take to contribute to the ECB process. One example is the leader of the organization, in this case, the ED who can help shape the vision for the evaluation and ensure data and communication flows smoothly. EDs also can help marshal the necessary resources and enable professional development for staff. They can be the drivers and nurture the organization's commitment to ECB in an active manner by understanding that their staff may have particular needs for understanding, tools, and support. Figure 23 below shows that the EDs were picked slightly more than evaluators and program staff as the most important stakeholders for ECB, and those that picked EDs described them as makers or breakers for the process.

Figure 23 – Most Important Stakeholders for ECB (According to Evaluators)



Some examples of responses help to convey why evaluators put EDs ahead of themselves for being most important for the ECB process.

- “I mean, I think the primary constituency for this is the leadership of the organization. The engagement of the ED could quite ultimately decide that attention needs to be diverted this way as opposed to that.”
- “Um, the more commitment you have from the top, the easier it is.”

- “I think it's critical that (pause) I mean, that's top leadership. I mean, those leaders have huge influence. And when a leader knows that evaluation is important, values it, finds budget resources for it, makes sure (pause) ensures that we're measuring progress and impact and outcomes, you know, it makes all the difference. I mean, it's like the parting of the Red Sea. It's like, you know, all of a sudden everything works. You know, people usher you into meetings and invite you in.”
- “Well, I think if you don't have the ED's involvement you might as well stop right away.

Evaluators encourage their peers to locate stakeholders lower in the hierarchy of the organization who have an interest in program evaluation can serve as valuable supporters for and participants in ECB. About half of the evaluators talked about how important it is to have the cooperation of program managers and program staff who they see as valuable resources and providers of critical information for the evaluation process. One evaluator mentioned she always reminds her clients' program staff up front that they are the experts for their programs.

The board of directors was referenced as important stakeholders by 12 of the 20 evaluators. One internal evaluator at a large youth-serving NHSO stated she would like to have a functioning board committee that's involved in the evaluation process. Another evaluator put the importance of the board's involvement into perspective:

“I think the boards of these organizations have to be on board and on track. And they have to for two reasons. One, they are supposedly guiding policy, setting policy, and so they should be well informed about what's going on and whether it's working and if it isn't, why. So they have to buy into it. Two, they also have to be involved in finding external resources if they need a third party evaluator.”

Funders were discussed for their important role as catalysts in demanding program evaluation from their funded NHSOs and providing them with an opportunity to recognize their external context and integrate those demands into their ECB efforts. They

also mentioned the more obvious important role they have in funding program evaluation. Financial resources are critical to ECB, and funders have the opportunity to help their grantees to build capacity to provide them with more meaningful information about the programs they fund. Funders can also be the facilitators of a more flexible evaluation process and a less tense atmosphere with a risk-friendly environment. One evaluator explicitly states her view on funders' role, "I think if they're going to require it, it's kind of incumbent on them to help build the capacity of those organizations that are going to come to them for money."

Evaluators were asked what their biggest obstacles were for building evaluation capacity and here were their responses in order of frequency:

Time constraints – 10

- Not enough staff time dedicated
- Not enough time allowed for evaluator

Culture – 9

- Lack of understanding its importance
- Undervaluing
- Lack of commitment
- Lack of support
- Resistance
- Not realizing benefits

Funding – 9

Staffing – 3

- Turnover
- Availability of human resources

Knowledge/expertise

Scheduling

Counterproductive accountability constraints

- Funder or accrediting body requirements are hindrances to evaluation

Lack of ownership

- Organization does not have dedicated evaluation function or point person

Multiple sites

Technology

- Organizations lack appropriate technology

Evaluators presented a mix of issues that are predominantly organizationally driven and specific tactics or methods evaluators have used that have been successful for ECB. For the organizational context, they referenced culture a total of 80 times in 20 interviews usually discussing how important it is for an organization to develop a culture that is supportive for evaluation and ECB. The most important aspect of an organization's culture and the most frequently referenced at a total of 45 times was commitment. Some adjectives they used to describe the value of commitment were "critical," "required," and "necessary."

Relationships are a critical part of ECB, especially in NHSOS who are in the people business. All evaluators described their relationships with stakeholders as one of the most important part of their job. Stakeholder involvement was repeatedly referenced throughout the interviews (54 times) as a required part of the ECB process. Evaluators offered some strategies, tactics or methods they have found to contribute to successful ECB endeavors, and the majority of them focus on stakeholder involvement. The evaluators paint a picture that indicates the days when they could get by with just being experts in the mechanics of their profession are long over. One commonly recommended tactic is to communicate in a language stakeholders can understand to reduce tension, insecurity and help to demystify evaluation for them. Several discussed at times not even using the word evaluation when beginning projects. They instead ask the stakeholders what they're looking for so they realize they're making decisions rather than an outside expert coming in to evaluate them and their program.

Another recommendation is to have empathy for what program staff have to deal with. In the world of NHSOs, they are typically over stressed and overworked. Showing

a little understanding and compassion for what they do and what it takes for them to do their jobs not only helps build the relationship but also conveys that the evaluator has included their insight and perspectives in the process of learning about the program. The evaluator can also begin to strategize on how to minimize the burden of program evaluation for the staff and value their time. Lastly it helps remind them that they are important to the ECB process. One evaluator explains, “Setting the stage and trying to build that relationship from the get-go is very important. Always reminding them that this is my responsibility, but you’re the expert in your program. I have expertise in statistics, and pulling data together and making that valuable to you. But you know your program so you have to be at the table.”

- The use of logic models and their role in ECB was discussed, and all 20 evaluators stated they utilize logic models in their program evaluation process and all affirmed that logic models play an important role in ECB. Here are examples of how some evaluators discussed the role of logic models in ECB:
 - “We’re not going to evaluate everything on our logic model, but we can do that, and we shouldn’t do that. But let’s be thoughtful and what do we look at, what do we learn, and how does the logic model need to change? And that it really feeds into the next year. It needs to be updated. So to me, I talk about a logic model as a program management tool.”
 - “Usually I’ve discovered is that it’s good to sit down with them and discuss their programs. I don’t believe in evaluating sort of I guess swooping in as the expert telling them what their program is. I have to have them talk through what their program is. Ideally I like to take people through logic model and looking at making what they’ve been doing be transparent.”
 - “I do think they contribute to evaluation capacity building. Because I think understanding that you’re going to evaluate what everybody agrees to be logic of the program really helps people find the utility of that. Really buy into and say okay, if this is the model, this is what we’re doing, this is the outcomes that we expect to get. Now let’s go see if you’re actually getting those outcomes.”

- “I think that it’s probably the best way to get people to think about the measures that they need to be collecting also to be thinking about in terms of what process information they should be collecting. I think one of the biggest challenges though and this is a place where capacity (pause) it’s very difficult to rally building capacity for data analysis.”

The profile information for the 20 evaluators is shown below in Table 15:

Table 15 – Evaluator Profile Information

MEAN *TENURE	MEDIAN *TENURE	RECEIVED INSTRUCTION** IN EVAL.	HIGHEST LEVEL OF EDUCATION
17.3 years	17 years	ALL – 30% C&W – 25% C – 40% EI & W – 5%	60% - Doctorate 40% - Master’s

n=20

*Tenure is total years served in profession

**Instruction Key: ALL = College, Evaluators Institute & Professional Dev. Workshops
C = College
EI = Evaluators Institute
W = Professional Dev. Workshops

The tenure for most evaluators places their beginning in this line of work in the 1990s when the demand for program evaluation significantly expanded in the nonprofit human services arena, largely due to the United Way and their outcomes measurement movement. As expected, all evaluators have had instruction in evaluation, with more than half receiving it in college. The higher levels of education for evaluators, as compared to the other stakeholder groups, are a reflection of the standards for academic requirements in the profession. For the most part, an education in evaluation can only be experienced at the master’s level.

Case Studies

Two purposeful case studies examined the NHSOs Families First and Decatur Cooperative Ministry utilizing Volkov and King's *Evaluation Capacity Building Checklist* (2007) and Stufflebeam's *Institutionalizing Evaluation Checklist* (2002), as a framework for analyzing their ECB efforts. Items in these checklists will be referenced in italics following the checklist authors' last name initials, i.e. (V&K) for Volkov and King and (S) for Stufflebeam, when they are represented by examples in the case studies. Families First, a large NHSO, and Decatur Cooperative Ministry, a small one, were intentionally chosen to serve as potential models for ECB based on stakeholder interviews, a review of the organization's documents and web sites, and the researcher's participation in various evaluation-related stakeholder meetings.

Families First

The mission of Families First, located in Atlanta, Georgia, is "To strengthen and preserve families in partnership with them and their communities" (Families First, 2007). The organization works toward achieving this mission through seven program areas including adoption, foster care, residential services, employee assistance, counseling services, divorce education, and child visitation. In 1988, the organization was closing in on its 100th anniversary, and with an annual budget of about \$1.5 million, decided to conduct their very first capital campaign to fund new programs. The programs were based on the latest research on service delivery for family-based social issues. The organization also began to formally build capacity for and engage in program evaluation, eight years before the United Way's outcome measurement movement and about twelve years before literature on ECB appeared in publications.

Bert Weaver, the Executive Director of Families First in 1988, realized he needed a way to demonstrate the effectiveness of these new programs to the donors who took the risk of funding them. At that time, funders were beginning to ask critical questions about programs they supported and their effectiveness. He decided the time had come for the organization to formalize its approach to program evaluation. Weaver reflects, “I knew that we needed to get very serious about outcomes in saying this program works or it doesn’t work. That’s really what kind of got us going in that direction” (personal communication, January 23, 2007). Weaver recognized both *the internal and external organizational contexts* (V&K) and *identified, supported and addressed internal and external driving forces* (S) for program evaluation. He decided to *personally support and share responsibility for ECB* (V&K) and *located another evaluation champion in the organization* (V&K) in Chris Valley.

Valley, who was then the organization’s Director of Program Development and Research, responded and sent a memo to Weaver on October 18, 1988 calling for a “professional approach to program evaluation” and a new, full-time program evaluation director position. Valley reasoned that this new person “...would add a new dimension to agency efforts in monitoring program performance, marketing services to funders, and the possible redesign of services. It would be an invaluable support to agency program development” (personal communication, February 15, 2007). Weaver and Valley set out to *increase the organization’s interest in and demand for evaluation information* (V&K). However, beyond their explicit support and support from Pat Showell, then Vice President for Programs and currently their ED, they had to determine if and *to what extent the internal environment was supportive of change* (V&K). Weaver presented the

organization's intended new direction and Valley's proposal to the board of directors. The board reacted positively to this new direction and voted to extract funds from the capital campaign to support the new position. In creating this new position, Families First was able to *assign responsibility for facilitating the ongoing development and evaluation of evaluation processes (V&K)* and begin to *engage and support a capable, credible evaluation team possessing expertise in field work, group process, interviewing, measurement, statistics, surveys...(S)*.

Hugh Potter, now with the CDC, occupied Families First's new position of Director for Evaluation. Potter started out evaluating programs for teenage pregnancy prevention, domestic violence intervention and children whose parents had divorced. He described the process in the beginning as challenging but progressive for its time and the fact it was a new initiative in an organization that had existed for 100 years:

"One of the things that were brilliant for me working there was when we would get a new idea, they involved me from the beginning. So, in many ways the environment I worked in at Families First was an evaluator's ideal. You were involved from the beginning of programs, drove the social workers crazy, and sat back and asked people what they wanted to achieve" (H. Potter, personal communication, February 7, 2007).

Potter described how he attempted to establish and apply *clear, appropriate evaluation purposes* such as *improvement, accountability, organizational learning and dissemination (S)* and *promote and facilitate people's learning evaluation by involving them in meaningful ways in evaluation planning and implementation (V&K)*. He discussed how he attempted to build program staff buy-in for evaluation through an active learning process. For example, he discussed the programs with staff to get them to sharpen the focus of an evaluation question:

“Saying, ok here is what you’re designing in this program. What do you really want to accomplish? We would sit down and figure out what it is that they considered a marker of success or a marker of failure. And then I would drive the social workers crazy by asking them how do you recognize it? And of course one of the things social workers are famous for saying is ‘I know it when I see it.’ And I would ask them how they knew and would just keep asking that question” (H. Potter, personal communication, February 7, 2007).

Potter confirmed that his strategies, along with the support from the organization’s leadership, helped him work to *increase the organization’s interest and demand for evaluation information* (V&K).

King and Volkov’s recommendation to “*allow adequate time and opportunities to collaborate on evaluation activities...*” (K&V, 2007) played into Potter’s efforts to ensure program evaluation evolved into being part of the work of the social workers.

“One of my experiences was that I knew case workers didn’t have extra time to fill out forms, so one of my goals was always to integrate program evaluation data, kind of what we do with surveillance data here (at the CDC), so it needs to be part of an ongoing process. It doesn’t need to be something you need to comb records through to answer a question for me the evaluator” (H. Potter, personal communication, February 7, 2007). Weaver adds, “For time, I would give that to the senior administration and the program directors, as they bought into the concept they knew they had to have time. You know that was an issue that was dealt with in terms of trying to minimize the amount of intrusion and on the other hand recognizing that there had to be some time committed to it. We had performance criteria for people to reach their benchmarks. So that was always taken into consideration, and there was time used for this, record keeping and various things that had to be done” (personal communication, January 23, 2007).

Potter was asked how in the beginning of the ECB process apprehension among program staff for program evaluation changed to interest and excitement.

“I don’t know if I can take any credit for this. I think you have to go back to the leadership in place at that time. One thing I will say, I was sort of part of junior management. But they really allowed me to come in and work with them and we did change management. I will say also some of

the members of the board were really helpful here” (H. Potter, personal communication, February 7, 2007).

Weaver helps describe his role, as leader of the organization, in this transformation.

“I had very strong feelings that this was very important and a component whose time was due. I wasn’t preaching and I wasn’t dictatorial about it. But I was convinced if we were going to survive as an organization, we needed to know if these programs were making a difference or not. So, I had a conviction about it and I think that the staff saw that conviction in me. The fact that I was willing to go out there and bust my butt to raise money for that, along with all the other things I was raising money for – to make the place livable and look decent, to develop some new programs, which staff wanted developed. I think all of that kind of was a message. So, I was very involved in raising the money for it and I was very involved in getting the board to buy into it. We worked with the management team for months, and we did it in steps. I think all that was convincing and reassuring, and I think the staff was seeing that this as a new direction we were going in. Everyone was going to be on board, and it wasn’t going to be for an elite few (personal communication, January 23, 2007).

Weaver also explained how important relationship building, especially between the evaluator and the program staff, was in contributing to this culture change.

“I think to have a researcher who respected the programs, and the program directors buy into that, was and still is a turning point factor. As opposed to a researcher who maybe was looking to see if there are problems. It shouldn’t be kind of ‘I gotcha.’ Then the staff are going to sabotage you every time, and why not, they’re protecting their backsides. What we did was a fair amount of education with staff to make them understand that we’re not here looking to see if you’re competent or not. We’re really seeking to see if we were making a difference or not” (B. Weaver, personal communication, January 23, 2007).

Families First established *a capable ECB oversight group (composed of members of the staff, board of directors and community to initiate, evaluate, and advance evaluation processes continually in the organization (V&K).* The group, now called the Evaluation Committee, has been a board committee since inception. It consists of the executive director, director of programs, director of evaluation, three board members and

outside evaluation experts, usually from local universities. The committee's purpose is "...to review, critique, comment, and recommend necessary changes in evaluation designs, procedures, and reports." Potter, Robert Fischer who came after him as the director of evaluation, and Peter Lyons, their current (external) evaluator and a professor from Georgia State University, all indicate the committee has been very helpful for their work. The collaborative efforts of these evaluators with the committee demonstrate Families First's effort to *use evaluation personnel effectively* (internal professionals and/or external consultants) (V&K).

Potter explains how he and some members of the advisory group worked to increase stakeholder support for program evaluation.

"We built a good advisory group. This guy who headed it up was a vice president down at the IBM building. And one of the things he and I worked on is what we might call building a business case. We began to show how evaluation was part of the process of planned change, and that we were really about improving client services. And so I think what helped us was that we sold it on two levels. We sold it to the board as something that would make us a stronger provider of services in the community. And we went to the people inside to show them how they could do their work even better than they could adapt to" (H. Potter, personal communication, February 7, 2007).

Valley confirms the board's role in the ECB process. "This board was one of the most incredibly risk taking boards. They stepped forward when no one in this community was doing outcomes research. And the board said, okay we'll do it" (C. Valley, personal communication, February 15, 2007).

Potter reinforces that this cultural change at Families First involved a variety of stakeholders at different levels and in different capacities.

"And let me say, this is where Chris came in so wonderfully. Because he was not only the development manager in those days but he was also more or less communications and all of that. We worked very much hand and

glove to put those sorts of things together and to sell this to people. So, it was this whole process of organizational management, planned change, clients' service improvement, etc. That's why during the time I was there and I think for the next couple of people who came after me, why it sold" (H. Potter, personal communication, February 7, 2007).

These efforts described by Weaver and Potter were part of Families First's effort to *promote and support stakeholders' buy-in, participation, and support from all levels ...to assure that top management and governance are knowledgeable, supportive, and involved inn the evaluation effort* (S).

Families First's capital campaign in 1988, in addition to funding the new programs and the new evaluation director position, also established a program endowment fund that funded, among other things, evaluation expenses from its interest and investment earnings. This served as a *revenue-generating strategy to support program evaluation, and assured long-term fiscal support and explicit, dedicated funding for program evaluation activities* (V&K). The endowment is currently has more than \$1 million, and it has *supplied the evaluation effort with sufficient funds, facilities, equipment, services, software, and technical support* (S). The endowment has enabled the organization to internally fund a myriad of evaluation projects, including some costly longitudinal studies, over the past 20 years.

Rob Fischer followed Hugh Potter as director of evaluation at Families First. Fischer *developed and utilized an internal reporting/monitoring/tracking system and developed an effective communication and reporting capability* (V&K) to relay evaluation findings to evaluation stakeholders. Weaver and Valley encouraged and supported Fischer's efforts to publish journal articles, and Families First evaluations are featured in *Research on Social Work Practice, Journal of Divorce and Remarriage, Child*

Welfare, and *Family and Conciliation Courts Review* among other publications. Fischer cites Weaver's leadership as the driving force behind enabling him to advance program evaluation in Families First. "He was vitally interested in evaluation, a real champion of it, and he asked tough questions."

The issue of managing the political environment for program evaluation has played out in the ECB progress at Families First. Weaver reflects on having an internal evaluator during his tenure, "The danger for an in-house person is that they are a colleague, they are a staff member, and they are a friend. And if they're not really professional, that sharpness can begin to be co-opted. So you want to be careful about that" (B. Weaver, personal communication, January 23, 2007). Weaver commented that Fischer was successful in balancing the importance of relationships with program staff while not letting them intrude on attempts to remain objective. Peter Lyons, an external evaluation consultant and professor from Georgia State University, discussed what made him decide to become the next evaluator after Fischer.

"I was only interested in doing it if they were really interested in genuine program evaluation. I didn't want to do something that says this was just a snow job. Chris' response was that I would report whatever comes out. And I've been able to say, you know, it isn't working" (P. Lyons, personal communication, January 24, 2007).

The documented ECB efforts of Families First contain themes related to organizational learning and several stakeholders provide their take on it. Valley explains, "All of our evaluation is a point of departure for debate, and it's about what can we learn about our programs" and he adds, "We learn every day, first, from our clients, second from our peers, and third Families First learns as an institution from our program evaluation because that gives us as an organization the opportunity to learn together" (C.

Valley, personal communication, February 15, 2007). Pat Showell, their current ED who succeeded Weaver, describes how program evaluation is driving organizational learning at Families First.

“I think we are more of a data driven organization. We come to issues, problems and opportunities out of that data, that knowledge, rather than our emotions. That’s a good thing. We’re a social work organization, so we all have our opinions and emotions. Evaluation provides a framework that we wouldn’t have otherwise” (P. Showell, personal communication, December 11, 2006).

Program evaluation appears to be embedded in the culture of Families First.

Potter, Fischer and Lyons all indicate they have received the proper support and resources to effectively do their jobs from either Weaver and/or Showell indicating that leadership succession has not impacted the organization’s valuing of program, evaluation. Showell confirmed her intent on advancing the ECB foundation left to her.

“There was a belief and vision that it was important not only for the stakeholders and funders, but it was also important for us. So, I see my role as picking up on the prior CEO’s belief that if we’re going to make a difference and have impact, we need to be able to know beyond the warm and fuzzy stories we have.” She describes the role of program evaluation as “...integral to our ability to provide the kinds of services that our constituents need as well as the quality of services” (P. Showell, personal communication, December 11, 2006).

Lyons describes program evaluation as “...such an integral part of what they do,” (P. Lyons, personal communication, January 24, 2007), and current Board Chair, Mary Yates calls it “...a tradition at Families First” (M. Yates, personal communication, January 30, 2007). Pat Pillow, the current Vice President for Programs, sums up Showell’s role in influencing the organization’s culture.

“I think the culture here understands that it’s part of doing business and it’s a necessary part of what we do. I think she’s the driving force behind all of that initiative. I think when she’s having the agency value it and

creating the resources to have it, it comes from her as the top down commitment” (P. Pillow, personal communication, February 12, 2007).

The last twenty years of Families First’s efforts in ECB demonstrate it’s an ongoing process, it has evolved and progressed over time, and it has provided benefits to the organization. The early ECB efforts of Families First, according to Potter, reaped future dividends for their evaluation and program delivery processes.

“After I came back and talked to Chris occasionally I could see the benefits of some of the things we started in the late 80’s. I think it’s had commercial benefits as well as client benefits. I think they provide better services and they’re better able to target their clients. But I would also add that they have expanded in size and scope, and they have developed what we would now called evidence-based products” (H. Potter, personal communication, February 7, 2007).

Potter summed up his experience at Families First.

“I really do look at what the leadership Chris, Bert, Pat and others did as incredibly visionary. It is something where I look around the nonprofit world that I work in now, between criminal justices and public health, and really don’t see many places that have that kind of vision. And I work with organizations from all around the country. In many regards I look back at what I got to do at Families First as almost an ideal situation” (personal communication, February 7, 2007).

Decatur Cooperative Ministry

Decatur Cooperative Ministries (DCM), located in Decatur, Georgia, provides a continuum of services to at-risk and homeless families. Their mission reads: “Decatur Cooperative Ministry serves our neighbors in need. Together, we strive to end homelessness, empower our community through education, and celebrate our faith-based diversity” (DCM, 2007). The organization, founded in 1969, has an annual budget of approximately \$500,000. The organization’s strategic plan for the period of 2003-2006 contained four primary goals, one concerning effective and efficient program services, so

DCM embarked on evaluating all three of its programs – Project Take Charge, Hagar’s House and Family Transitional Housing.

Beth Vann became DCM’s Executive Director in 2001. The organization already had *a capable evaluation oversight group* (V&K) in place with the Organizational and Program Assessment work group comprised of board members, staff and volunteers. The charge for the group is to “Conduct comprehensive assessments and evaluations of DCM programs, services and organizational effectiveness; develop a process and a schedule to evaluate various programs; present written reports and recommendations to the board and staff for proposed areas of improvement.” The committee was developed by board member Sarah Gill who worked for a large evaluation consulting firm in the Atlanta area. Sarah was able to *use evaluation personnel effectively* (V&K) when she procured *a capable, credible evaluation team possessing expertise...(S)* comprising professional volunteers including two evaluators at the CDC and several graduate students from the Rollins School of Public Health at Emory University.

Vann notes that Sarah “...was willing to really drive the process where I could be part of the process, but it wasn't that I had to drive it” (B. Vann, personal communication, December 18, 2006). Vann admits if she had to be the driver of the program evaluation process, it likely would not have happened because of her responsibilities of managing crises, being short staffed, and keeping the organization financially healthy. She also attributes the board’s interest in critically looking at the programs as a driving force, allowing her to “...want to look at the programs harder, as well as look at the whole organization and its structure” (B. Vann, personal communication, December 18, 2006).

Vann did, however, personally *support and share the responsibility for ECB* by participating in evaluation team meetings and ensuring program staff participated in the process and worked with the evaluators. Vann *established clear expectations for people's evaluation roles and provided sufficient time during the work day for evaluation activities (V&K)* and *promoted and supported stakeholders' buy-in, participation, and support (S)*. She did this by meeting with the program staff as a group and individually to 1) explain why the programs were being evaluated; 2) ensure they understood that having quality evaluations conducted by qualified evaluators volunteering their time was an opportunity the organization could not pass up; and 3) emphasize how important their cooperation and participation in the process, especially for working with the evaluators, was critical to the success of the evaluations. Vann did this to get a pulse of how the staff would embrace the evaluations and *determine if and to what extent the internal environment was supportive of change (V&K)*. This was a potential concern because this was the first time that DCM's programs were formally evaluated and they had a new manager of one of their programs. Vann summarized her efforts, "So it was just getting everybody on board and making the time and allowing the time for that" (personal communication, December 18, 2006). She modeled behavior by personally attending evaluation team meetings and helping them coordinate the evaluation process. Vann added, "It was really just letting the staff know, getting buy in, hearing what their thoughts were, and creating the time, access and space to do it" (personal communication, December 18, 2006).

The program staff's buy-in and commitment for program evaluation was conveyed in their interviews. For example, Cliff Richards, the Project Take Charge

Program Manager described the importance of evaluating programs as inherent to satisfying the organization's mission.

“So I think that in doing a program assessment you have to not only look at what you're doing right now, but also see if you are fulfilling that mission. You have to determine if there are ways for you to make some changes to meet that mission and vision more specifically and help the program be more fulfilled. And I say the word ‘fulfilled’ because growing bigger does not necessarily mean success. And DCM, I think, is very conscious of that. We're not interested in being the Salvation Army or the St. Vincent de Paul's Society. We feel more effective in the size that we are” (C. Richards, personal communication, February 6, 2007).

Sybil Corbin, Hagar's House Program Manager, stated that she thought DCM engaged in evaluating their programs because “You would want to know at some point the effectiveness of what you're doing and have a way to determine that and measure that, or perhaps a more professional way to assess and determine that” (S. Corbin, personal communication, January 23, 2007).

The commitment to program evaluation from program staff was also demonstrated by their actions. Vann's efforts to prioritize program evaluation and make sure staff dedicated a portion of their time allowed *opportunities for sufficient input in decision making for the programs* (V&K). Program staff sat in on that initial planning meeting and was interviewed by the evaluation team. They provided the team with access to files. They briefed their program participants on the project and let them know evaluators would be contacting them. They also played primary roles in implementing the evaluation recommendations, some that have been completed and some that are still in progress. Their participation, Vann's support, and the evaluation's team framework established the *incorporation of a feedback mechanism in the decision-making process and an effective communication system* (V&K).

Vann recognized external and internal reasons for evaluating DCM's programs.

“We invested a lot of time...in doing the evaluations and we also had a work group that invested time implementing the recommendations. It's given us things to really help bring our programs up a level, run better, more smoothly. You know, everybody wants a logic model these days and has their format. But the thing is they did these really great logic models. So I feel like we know a lot about that and have the capacity to create that kind of thing now. When they ask about evaluation in grant applications or in a site visit, we can really clearly point out what we've done. I think it gives us really strong positioning to talk to granters that we really do take all this seriously. And that we do really look at ourselves and want to do the best we can” (B.Vann, personal communication, December 18, 2006).

The collaborative and participatory framework for program evaluation established by the evaluation team and Vann involved staff, board and volunteers. This helped *promote and facilitate people's learning evaluation by involving them in meaningful ways* (V&K). Vann reflected on some examples,

“As I think of Project Take Charge, it really changed the way that the program manager thought about the financial management classes, for example with the materials and the curriculum. Some of what came out in that evaluation for Family Transitional Housing is the need for better training of, and more involvement from, the volunteers so they feel more connected. And so now that's something we're working on for that particular program. I think overall, it just reinforced the importance of their work and the program.”

DCM with Vann's leadership, board involvement and the efforts from the evaluation team has all contributed to ECB within the organization. However, qualified volunteer evaluators may not be available in the future, and the organization did not *assure long-term fiscal support and/or come up with revenue-generating strategies to support ECB* (V&K). DCM will have to strategize how in the future they can *supply the evaluation with sufficient funds, facilities, equipment, services, software and technical support* (S). Strong commitment exists from Vann, her board and staff for continuing to

have programs evaluated which presents the organization with the challenge of ensuring that commitment can be realized in the future.

Conclusions and Recommendations

The following conclusions and corresponding recommendations are drawn from cross analyzing all 126 interviews as a total group of participants, within each stakeholder group and the case studies.

Leadership from the executive director is critical for successful ECB

EDs are in an important position to influence their organizations' efforts in ECB. The EDs who took all or almost all eight steps for ECB (see Figure 15) as well as the two EDs of the case study organizations lead Type I organizations. All of these leaders strongly influenced, developed and shaped a culture for continuous improvement in their organizations. While almost all EDs made the intellectual connection of program evaluation to their organization's mission, only those who used this connection as a value orientation, and in some cases also as a management tool, were successful at ECB.

EDs have at times strived to build capacity for their organizations to have more facilities, acquire more resources, and serve more consumers. Can they also build capacity to evaluate and demonstrate the effectiveness of their programs? Can they work towards embedding program evaluation into their organization's culture to where it becomes a basic assumption and natural part of their work? More evaluators in this study, 35%, chose EDs as the most important stakeholder in the ECB process than any other stakeholder group. They explained that without top down support from the organization's ED, the culture, resources and structures for ECB will likely not exist. To support their point, a NHSO can have funds, personnel, expertise, and data collection systems but if the ED does not prioritize, plan, and budget for it the organization will not

likely engage in ECB. We therefore can hypothesize that without proper leadership from an ED a NHSO is unlikely to successfully endeavor in ECB.

The Board's role in ECB is underutilized

The results of this study show that the board's role in ECB for their organizations is not being used to maximize their contributions. There are several factors that appear to be hindering this objective. First, the content of the ED, board chair and program staff interviews conveys that all three stakeholder groups at times intentionally or unintentionally develop a mindset in their organization that believes policy and programs, and those responsible for each shall remain separated. One example was that only about 9% of the board chairs thought ensuring programs were effective was a part of the board's accountability while 30% believed it was a part of the organization's accountability.

Second, the majority of EDs, board chairs and program staff confirmed their boards primarily focused on fiscal matters. Approximately 27% of the board chairs indicated fiscal matters were a part of the board's accountability making it the most frequent response. Financial stability was the board chairs' most frequent response for their top, second and third priorities, while only one board chair included program effectiveness as a priority.

Third, in spite of the fact that 95% of the board chairs explicitly connected program evaluation to their organization's mission, similar to the EDs efforts to cascade that value orientation, only those board chairs who took an active role in ECB, beyond just reviewing information, lead Type I organizations. The two organizations in the case studies as well as the majority of the Type I organizations had working evaluation

committees on their board. Fourth, the responses from the majority of the board chairs' on their perceptions for program evaluation demonstrated a lack of understanding for the concept and process. All of the Type I board chairs had some instruction in evaluation while only about half of the Type II board chairs had such instruction. Until board members become educated on program evaluation and realize its importance as a policy and management tool, they are likely to not engage much in ECB and leave it up to their ED and staff.

Program staff's internalization of program evaluation presents challenges

Human service professionals are generally known for their intensity, passion for their work and caring for their consumers. While these characteristics are admirable, necessary and should be expected from the program staff of a NHSO, they can be counterproductive for effective ECB. Program staff through this intense focus may internalize their perspective of program evaluation to where they may misunderstand the concept, not realize how important their role is in program evaluation, and/or have the notion that management or the outside expert evaluator is responsible for it. Several examples were evident in their descriptions of program evaluation where they more so described their job of delivering the program. Those that described it as determining if goals or objectives have been met did so in relation to meeting the goals or objectives of their specific responsibilities, tasks and functions as we would expect in employee performance as opposed to the larger context for the program. This was also reinforced by the program staff that characterized it as a process for solely assessing consumer satisfaction, which is an important part of program evaluation but may not get at whether the consumers experienced the intended outcomes of the program.

This internalization can be also counterproductive to ECB, but in a different direction. The most noticeable examples were found in all the psychologists, therapists and counselors who participated in this study who described program evaluation as an administrative function management was responsible for and one they did not participate in. They gave these perceptions even though they directly provide the program's intervention to their consumers; and they conduct evaluative functions such as evidence-based practices, reflective learning, consumer pre and post assessments, tracking of consumer progress, and revamp their method based on the latest research and observing what works best in their practice. Their misperceptions were massaged a bit when particular program staff were probed further in the interviews. They were admittedly unexpected given the explicit presence of evaluation in the ethical and professional standards for the social work, psychology, counseling and workforce development professions.

It appears that this internalization also impacted the fact that only one program staff interview participant, from a Type I organization, described her role as being an important resource for evaluators by providing information about their work and the program. This is an important point for ECB that relies on stakeholder involvement, especially from the front line workers delivering the program. One fifth of the evaluators stated that program staff are the most important stakeholder for ECB, and those who did not discussed at length throughout the interviews the critical nature of the program staff-evaluator relationship. Other primary stakeholders invested in ECB must strategize on how to break down these misperceptions for program evaluation from program staff possibly through professional development, networking and internal accountabilities.

Funders have an opportunity to expand their impact on ECB

The lack of cohesiveness from the 24 funder descriptions of program evaluation also reflected misunderstandings and misperceptions. Some thought it meant assessing the impact of their grants, some thought it was seeing how their money was spent and others thought it meant evaluating the entire NHSO. It may be no surprise that the three descriptions coming closest to a “textbook” definition of program evaluation came from three professional evaluators who had Ph.D.s and experienced evaluation in college.

About two thirds of the funders stated they ask their grantees for program evaluation information; however what they ask for is likely based on what they think program evaluation is. Funders have an opportunity to increase their knowledge and understanding of program evaluation so they can better understand what they’re asking their grantees to provide, what it takes to deliver and evaluate programs, and how they can better realize their potential role in ECB. Only 38% of the funders in this study that asked for program evaluation information helped pay for the evaluations in spite of 91% of the receiving some instruction in evaluation. This can send mixed messages to grantees and appear contradictory.

Only 20% of these funders thought that their similar type funders, for example such as other community foundations or United Ways, adequately financially supported program evaluation. This appears to be an issue funders are aware of, so it begs the question why they don’t provide more funds to help pay for something they say is important and require from their grantees. Two theories come to mind. One, they may be treating evaluation costs as administrative expenses rather than program expenses possibly with their grantees also treating evaluation costs this way. Two, funders may

possibly want NHSOs to build their own internally supported evaluation capacity to truly value it by investing their own funds in it. Only through such an understanding will they be able to properly value it, encourage it, model it in their own organizations and financially support it. There are resources and organizations such as the American Evaluation Association, AEA's Nonprofit and Foundation Topical Interest Group, local evaluation associations, colleges and universities, the Council on Foundations, the Foundation Center, and peer funders already substantially supporting evaluation among others that funders can turn to for professional development and networking opportunities.

An important benefit of a more thorough understanding of program evaluation is the opportunity for funders to have dialogue with their grantees. Of the funders in this study that didn't fund evaluation costs at all also stated they would consider funding them if grantees asked for them or included those expenses in a grant application. However they also stated that no one ever asks for it indicating that NHSOs may be contributing to this lack of financial support. Funders will most likely have to be more proactive and initiate dialogue with grantees over this issue and to learn more about what it takes to deliver and evaluate the programs they fund. This dialogue might help reduce the inherent tension of such an imbalanced relationship. Funders may consider thoughtfully and comprehensively discussing internally why they're asking for program evaluation information, what they're going to do with it, and incorporate feedback loops for communicating with their grantees. Without demonstrating that they actually read reports and do something with the information, NHSOs may continue to treat program

evaluation as an externally driven process from stakeholders who don't truly value it themselves.

Another issue for funders and NHSOs that dialogue might help address is the great variance among the information they require of their grantees. This presents challenges for NHSOs that most EDs and some program staff and board chairs referenced in their interviews. NHSOs are attempting to diversify their funding sources due to the turbulent and uncertain funding environment. Complex and at times competing demands from multiple funders can present challenges for NHSOs to satisfy them. Funders have an opportunity to attempt to reduce this complexity by furthering streamlining their reporting processes and through peer learning to see what successful practices exist in the sector. The evaluation personnel of a United Way in a large city and that state's department of community affairs realized their lists of grantees were almost exactly the same. They have embarked on reviewing their evaluation reporting requirements in order to streamline them, have their requirements be as similar as possible in order to make the reporting process easier for their grantees. This progressive attempt at reducing the complex challenge of meeting multiple external demands should be considered for replication in other large metro areas.

Evaluators' roles are expanding

Evaluation capacity has grown from the narrow perspective of skills, expertise and funds to a concept that considers many more factors, some covered in this study. This requires evaluators to expand their knowledge horizons and their toolkit of skill sets necessary to build ECB and effectively evaluate programs. Evaluators now have to be more conscious of contextual and cultural factors that will greatly affect the ECB process while

at the same time realizing their limitations in terms of what they can control and what they can't.

The momentum of the ECB concept presents critical questions for evaluators. What interpersonal and communication skills will they need to develop the relationships with others that are important for stakeholder involvement and ECB? Can they meet the challenge of working with critical stakeholders such as EDs and board members without getting co-opted? Can they provide professional development, peer learning or collaborative learning opportunities for nonprofit organizations expressing interest in ECB? How do evaluators, especially consultants, balance engaging in ECB activity with potentially working themselves out of a job? These are just a few issues evaluators face as the complexity of ECB is being realized, discussed, published, and enacted. Attempting to resolve these issues will better equip evaluators to carry out what some feel is their purpose, to improve programs and contribute to social betterment.

Education matters

The EDs in this study with higher levels of education tended to lead Type I organizations, especially those who experienced evaluation in college. Type I board chairs included Ph.D.s, and a higher percentage of master's degrees than Type II organizations and no high school graduates compared to 7% of Type II board chairs. The program staff of Type I organizations had a higher percentage of Ph.D.s, 29% more Master's degrees and no high school graduates compared to 29% of the Type II program staff. The most highly educated funders had a more comprehensive understanding of program evaluation and were the ones that funded it for their grantees. Evaluators in this study comprised 60% Ph.D.s and 40% at the master's level indicating a potential

imbalance in knowledge and skills among stakeholders that might impact the ECB process.

NHSOs might consider encouraging higher education opportunities for their staff as well as professional development opportunities that might better equip them to engage in meaningful program evaluation and ECB. The 100+ graduate programs in nonprofit studies or nonprofit management across the U.S. might consider incorporating evaluation courses for their students so if and when they become the future EDs of organizations they will be equipped to manage people, finances, market their programs, raise money for their programs and demonstrate the effectiveness of their programs.

Logic models are not being properly exploited as an effective tool in ECB

Logic models for programs have been a part of program evaluation for many decades. This study hints that they currently are not used to their potential in ECB based on the incongruent responses from evaluators, funders and EDs. All evaluators in this study confirmed they use logic models in their program evaluations and that they are a valuable tool for ECB. They described the role of logic models primarily as tools for bringing stakeholders together and engaging them in discussion about how the program works, what they intend the program to accomplish and how the program specifically will accomplish it – the program’s theory of intervention. A few evaluators reminded us that their purpose is not only for new programs and provided anecdotal examples of how they have used them with existing programs that were running for years, but their organizations had never established the theory of change. The advantages in these situations were a renewed and refreshed look at their programs and a renewed value for their work and their role in their consumers experiencing the intended outcomes.

Some evaluators stated that a logic model is also a program management tool and the foundation from which the data collection and measurement framework is established. Once the theory of change, intended outcomes and their indicators are established the discussion can begin for what kind of data are needed for those indicators and how the organization will collect that data, a critical component of ECB. The overarching themes from the use of logic models that came out of these discussions were ones that contribute to the ECB process such as stakeholder involvement, organizational learning, empowerment, accountability, mission-driven, and help in diffusing of political issues inherent in program evaluation.

Slightly more than half the funders that ask for program evaluation information indicated that logic models are part of that information. The majority of those that did not ask for logic models indicated they used to but they received push back from grantees. Some stated that their grantees never understood them or saw them as an administrative nuisance they were requiring of them. Most indicated their grantees did not understand that they were a valuable tool for them and not just an external demand from funders. A few of these funders admitted to their grantees they stopped asking for them because they didn't do anything with them once they received them. These scenarios specifically seemingly would reinforce their grantees' notion that they are just administrative requirements.

The funders' comments for the use of logic models support the comments from the EDs. More than half of the 42 EDs did not know what a logic model was. The others that did typically described them as an administrative component to program evaluation and a demand of external stakeholders. Some common refrains were, "Yes we have logic

models because the United Way requires them” and “I’m certain we have one, but I have never seen it.” We can conclude that these EDs were either unfamiliar with logic models, did not understand them, or did not appreciate their potential as a vehicle for establishing their program’s theory, enabling their organization to better explain how their programs impact their consumers, and engaging stakeholders for the program evaluation and ECB processes.

The value in logic models is not necessarily in the visual shapes, arrows and flow of the model or as a product i.e. a model on a piece of paper or on the computer screen. The value lies in their ability to engage stakeholders in the evaluation process by having them discuss their programs, establish program theory, and determine what they desire and expect their programs to achieve. Here are a few possible solutions stakeholders can use to work towards consistency in the understanding, valuing and use of logic models so NHSOs can fully take advantage of their value:

- Ensure EDs, funders, program staff, board members and other key stakeholders are involved in the discussions about the program and the model development or revision process.
- Change the perspective for logic models being external, administrative demands to powerful tools available to NHSOs to use for establishing their program theory, their outcome measurement framework and the foundation for them to use when explaining how they expect their program to impact their consumers.
- Educate stakeholders that logic models are not static and need to be revised to reflect changes in programs and how they address changes in the environment and their consumer’s needs.
- Use logic models to revitalize program staff’s views of their programs, their work and their role in impacting the lives of their consumers. This powerful paradigm shift can move them from viewing their job as a function or task to viewing it as an important component of providing the intervention and contributing to the organization’s mission.
- Use logic models as visual tools to explain programs to new board members, employees and volunteers so they may have a working knowledge of the programs and their expected outcomes.

- Use logic models in case statements for fundraising, so funders can have a better understanding of how the organization will address the particular social problem it desires to treat or solve.

Time – ECB's greatest challenge?

Time was the most frequently cited challenge for program evaluation by board chairs and program staff. Half the program staff at Type I and 43% working at Type II organizations stated that program evaluation presented trade-off costs for their time. Evaluators (30%) referenced it in their descriptions of program evaluation and it was the most frequently referenced as their greatest challenge for ECB, slightly ahead of culture and funding. They indicated this challenge is two-fold, the lack of dedicated time from key stakeholders, especially program staff, and the lack of adequate time for a program's cycle and the evaluation process. EDs can play a part in the staff issue by incorporating time spent with the evaluation process as part of their job description, accountabilities and performance appraisals. They can also play a big part in prioritizing program evaluation in their organization and reinforcing the dedication of time for the process. Funders can help the program cycle and evaluation process issue by acquiring a better understanding of the necessary time for both and how that weaves into their funding cycle. They can make adjustments in their timelines and reporting requirements to more realistically accommodate the required time for these processes.

Evaluators repeatedly mentioned how the lack of time for program evaluation can negatively impact the integrity of the data, stakeholder support and buy-in, and the overall quality of the evaluation. Lastly program staff can have an important role when developing a new program or adjusting an existing one by conducting time studies for key program functions as well as working with the evaluator to establish the program's

theory of change and how much time must the consumer experience it to experience the intended outcomes.

Implications for Practice and Further Research

Caution for the generalization of this study is reiterated here due to the small geographic representation and sample sizes. However, the results and conclusions can provide stakeholders invested in the evaluation of nonprofit human service programs with broader perspectives for ECB and program evaluation. They can also gain insight to what the other stakeholder groups might be thinking, feeling and what their needs, constraints, and assets for both processes are.

Executive directors (EDs) can gain a broader perspective on ECB and use such information to interact and work with their staff and board for ECB efforts. They can have a better understanding of how influential they, as the top paid and hierarchical staff, can be in the evaluation capacity building process. EDs can also be better equipped to develop a culture for continuous improvement in their organization and environment which enables and rewards organizational learning. Lastly, these leaders can learn how effectively managing and balancing the *external pull* of demands from funders, government agencies and accrediting organizations with the *internal push* of building the intrinsic motivation to evaluate programs within their organization will contribute to successful evaluation capacity building.

Board chairs and presidents can have a new or renewed understanding of program evaluation and what is required to effectively engage in it. Through this understanding, program evaluation can become a priority for boards that tend to focus on financial, governance, and policy issues while categorizing program evaluation as an operations matter better left to staff. Boards can enhance their relationships with their EDs through working together to build evaluation capacity including, but not limited to, incorporating

it into strategic planning, marshalling the resources necessary for it and impacting the organization's culture for prioritizing and using evaluation. They can also use program evaluation as a means to keep the board and other stakeholders focused on the organization's mission and consumers they serve.

Program staff are increasingly being recognized as key stakeholders in the program evaluation process. This presents them with professional development and learning opportunities where they not only can gain a working knowledge of evaluation to effectively participate, but also renew their value for the importance of their work. This study showed that the major issues for program staff were not fear or resistance as one might think, but a lack of comprehensive understanding of program evaluation, lack of understanding their role in the process, and a lack of time for them to engage in the process. Program staff can remind themselves that they are the experts for their programs and this expertise lends well to the program evaluation and ECB processes. Through this realization they can begin to strategize for the knowledge and skills they need, as well as how to streamline work if possible to free up time for evaluation.

The literature indicates that funders are increasingly requiring evaluation information on the programs they fund. Variance exists among different types of funders and even within each type concerning why they ask for evaluation information, what information is required, who pays for the evaluation, how the information is to be reported, and how it is to be used. Providing information that can educate and enhance the awareness of funders on evaluation capacity can help to narrow such variance and increase the level of funders' understanding and support of evaluation capacity. Another byproduct can be an enhancement of funder-grantee relationships where funders see such

relationships more as partnerships and the programs they fund more as investments. Funders may also see the costs for evaluation as an investment in the programs they fund and more as a cost of doing business, and a program expense, rather than an administrative expense. This can help enact a paradigm shift from nonprofit organizations conducting evaluations to satisfy the demands of funders to where they are conducted to acquire meaningful information about programs that is used to improve programs and that both the organization and the funder have a genuine, mutual interest in.

Evaluators have been steadily moving the issue of evaluation capacity building forward since the late 1990's and it is currently garnering significant attention and momentum. There is a new topical interest group (TIG) within the American Evaluation Association entitled *Organizational Learning & Evaluation Capacity Building* that in only a few years of existence already has over 500 members. The number of conference sessions on evaluation capacity building has also substantially grown over recent years. While the literature has focused primarily on what the evaluator can do to help build evaluation capacity, this study helps to enhance the evaluators' understanding of their limitations for and the impact of other stakeholders on ECB. Through this greater understanding, evaluators can more effectively interact with the key stakeholders invested in program evaluation in an effort to maximize ECB, while understanding what they can control and what they cannot. Evaluators can go beyond focusing on the skill level of their clients, typically program staff, to focus on leadership, organizational culture, structures and other aspects of ECB.

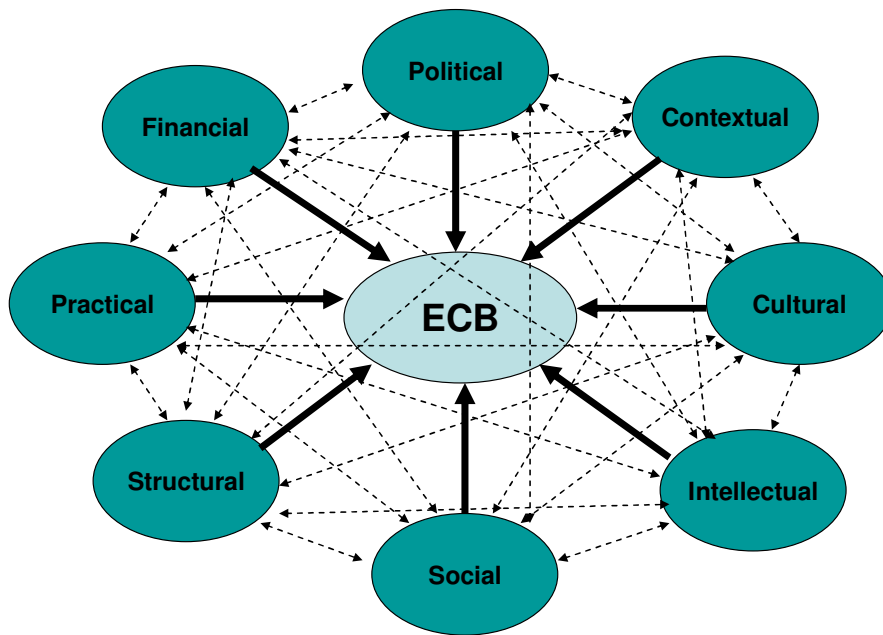
Scholars, researchers and practitioners have been discussing for more than a decade how nonprofit organizations are experiencing more attention, focus and requirements for the evaluation of their programs. This trend has particularly put pressure on nonprofit human service organizations due to the added dimension of the devolution of government. Motivations behind this movement may be driven by the increasing call for accountability, increasing competition for resources, quality signaling, public relations, a mission-driven sincere desire to effectively serve consumers, social justice or a mix of these reasons. This study does not intend to parse out these motivating factors to suggest that stakeholders need to further address why they want programs evaluated, as this is an area that warrants further study.

While this study does suggest that multiple benefits can be realized from program evaluation and that an organization's mission, covenant with its consumers and public trust would be the most effective intrinsic motivations likely leading to long term evaluation capacity building, it's primary purpose is to provide further insight into the dynamics and requirements for effective and efficient evaluation capacity building. When anyone wants something, their motivation notwithstanding, the next logical question they ask themselves is "What do I have to do to get this thing that I want?" If stakeholders truly want meaningful, practical, useable and feasible evaluations of programs they collectively must answer the question for what it will take to enable it and acquire it.

This study demonstrates that evaluation capacity building is an emerging and complex topic that requires efforts from multiple stakeholders. It has political, financial, social, intellectual, practical, structural, cultural and contextual considerations and

implications that also impact each other, as shown below in Figure 24. Stakeholders should recognize, understand, and factor them into their efforts to move forward.

Figure 24 – Considerations and Implications for ECB



With a comprehensive understanding of this complexity and the importance of their roles in the process, an expanded dialogue between them, and additional research for this topic these stakeholders can move closer to answering the question for what it takes to evaluate nonprofit human service programs.

Appendix A

March 19, 2008

Jill Jones
Anna Williams Children's Center
654 Boiling Springs Trail
Woodstock, GA 30189

Dear Ms. Jones:

I am a PhD student in Philanthropic Studies at Indiana University, living in the Atlanta, Georgia area, and I am conducting my dissertation research on the topic of program evaluation within human service nonprofits. You are one of 20 program staff randomly selected from the metro Atlanta and Indianapolis areas to participate in my study. The study includes face-to-face interviews that consist of 20 questions, take approximately one hour, and are intended to acquire the perspectives of program staff like you regarding program evaluation and issues of evaluation capacity. Interviews will be recorded and transcribed for accuracy; however strict *confidentiality* will be maintained with your responses as in accordance with my approval from the Indiana University Internal Review Board (IRB).

I believe this research is important and may be of interest to you because there has been little research that examines this topic from a multi-stakeholder perspective. I have also interviewed executive directors, board chairs, funders and evaluators. For example, I interviewed Karen Williams back in July, 2005 when she was executive director. I will be happy to share the results of the program staff study with you when the report is completed this spring.

I'll be contacting you soon to see if we can select a date and time that works with your schedule in the upcoming weeks. If you have any questions about this study, please contact me at (404) 297-9105 or salaimo@iupui.edu. Thank you for your consideration and I look forward to speaking to you.

Sincerely,

Sal Alaimo, MS, CVA

Appendix B

March 19, 2008

Ann Johnson
Evaluation Specialists
1042 Ship bottom Dr
Atlanta, GA 30303

Dear Ms. Johnson:

I am a current member of the American Evaluation Association and our Atlanta affiliate, the Atlanta Area Evaluation Association. I am also currently a PhD student in Philanthropic Studies at Indiana University, living in Decatur, Georgia conducting my dissertation research on the topic of program evaluation. You are one of 20 evaluators randomly selected from the metro Atlanta and Indianapolis areas to participate in my study. The study includes face-to-face interviews with evaluators of nonprofit, human service programs. They consist of 20 questions, take approximately one hour, and are intended to learn the perspectives of evaluators like you regarding program evaluation and issues of evaluation capacity. Interviews will be recorded and transcribed for accuracy; however strict *confidentiality* will be maintained with your responses as in accordance with my approval from the Indiana University Internal Review Board (IRB).

I believe this research is important and may be of interest to you because there has been little research that examines this topic from a multi-stakeholder perspective (I'm also interviewing executive directors, board chairs and funders). I will be happy to share the results of the evaluator study with you when the report is completed in early 2008.

Our interview is confirmed for Wednesday, April 16, 10:00am, at your office. If you have any questions about this study, please contact me at (404) 297-9105 or salaimo@iupui.edu. Thank you for your generous time and cooperation, and I look forward to speaking to you.

Sincerely,

Sal Alaimo, MS, CVA

Appendix C

Interview Instrument for CEOs of Nonprofit Human Service Organizations

1. What comes to mind when you hear the term *program evaluation*?

2. Have there been efforts made to evaluate your program/s?

IF EFFORT MADE – CONTINUE QUESTIONS 2-10

IF NO EFFORT MADE - SKIP TO QUESTIONS 7-9

(If yes - Probes)

- a. Please describe what has been done. When?
 - b. What methods were used?
 - c. Who conducted the evaluation?
 - a. Who selected the evaluator?
 - d. What was your role in the overall process from start to finish?
 - e. Did staff have a role in the evaluation process?
 - f. Did volunteers have a role in the evaluation process?
 - g. Did consumers have a role in the evaluation process?
3. Have the results of this evaluation been publicized internally and/or externally?

(If yes - Probes)

- a. In what mediums (newsletter, annual report, web site)?
- b. Was there any response to the publishing of the results? If so, please describe it.
- c. Who was the response from?

(If no - Probes)

- a. Are there any plans to publish the results?
 - b. In what venues and through which mediums?
4. Have the results of the evaluation been used internally and/or externally?

(If yes - Probes)

- a. Beyond publication

5. How much did the evaluation cost?

(Probe)

- a. Did the cost include staff time allocated for the evaluation?

6. Who funded the evaluation?

7. Was there a line item in your budget for program evaluation?

(If yes - Probes)

- a. Who decided to budget for the evaluation?
- b. Was evaluation budgeted for in advance of the evaluation?
- c. Was the budget for evaluation a one time occurrence or does your overall budget have an ongoing line item in it for evaluation?

(If no - Probe)

- a. Are there plans to budget for program evaluation? Regularly or annually?

8. Does your program/s have a logic model?

9. Does your organization have a current strategic plan?

(If yes - Probes)

- a. Is program evaluation included in that plan?

- b. If so, who suggested it be in the plan? Why?

(If no - Probe)

- a. Are there plans to develop a strategic plan?

10. Does your organization have any current contracts with government? (i.e., Federal, State, Local or with other nonprofits with which you are contracting and the source of the program funding is public/governmental funding).

(If yes - Probes) (If no – skip to Question 11)

- a. What are the terms of the contract (i.e., what program/service, over what time frame, with which client group, \$\$ amount of the contract)?

- b. Does the government contract and public agency require some form of program evaluation as a component of the contract?

- c. If yes, what is your organization required to submit as evidence of program evaluation?

- d. How does the government agency use your organization's program evaluation report? (i.e., funding decisions [contract continuation/termination], programmatic content, client focus, intervention modality, etc.)

- e. To what extent, do you receive feedback, technical assistance, or monitoring as a result of the program evaluation?

IF NO EFFORT MADE, SKIP TO QUESTIONS 12-14

11. Do you feel the evaluation effort was beneficial to your organization?

(Probes)

- a. If yes, why?

- a. What were the most important benefits for you as executive director?

- b. If no, why not?

IF EFFORT MADE, SKIP TO QUESTIONS 13-14

12. Could you discuss some of the reasons why your programs have not been evaluated to date?

(Probes)

a. Resources – staff, expertise, funding, time?

13. Do you have any future plans to evaluate your program/s?

(If yes - Probe)

a. Do you expect program evaluation to benefit your organization in any way?

14. Is program evaluation related to your mission?

(If yes/no - Probes)

a. If so, how?

b. If not, why not?

15. Tenure as ED/CEO

16. Top 3 priorities as ED/CEO

17. # of contracts with a governmental contracting agency

_____ Federal _____ State _____ County

Service area of Contracts _____

18. Have you received any instruction in evaluation?

(If yes - Probe) – Workshops, seminars, college courses?

19. Highest Level of Education High School _____ Some College _____

Bachelor's Degree _____ Some graduate School _____ Master's

Degree _____

Professional Degree (MD, JD, MPA, MBA, MSW) _____ Ph.D. _____

20. Major area of study in college or graduate school:

Public Administration _____ Social Work _____ Law _____

Business _____ Education _____ Social Sciences _____ Humanities

_____ Other (specify) _____

Appendix D

Interview Instrument for Board Chairs of Nonprofit Human Service Organizations

Name: _____ Organization: _____

1. What comes to mind when you hear the term *program evaluation*?

2. What does accountability mean for the board as the governing body?

3. What does accountability mean for the organization?

4. Has program evaluation ever been discussed at board meetings?

(If yes - probes)
How often has it been discussed?

Who initiated it being discussed at the meeting (s)?

Can you describe what was discussed and why?

5. Is program evaluation designated to a particular committee or individual Board member?

6. Have there been efforts made to evaluate your program/s?

IF EFFORT MADE – CONTINUE QUESTIONS 6a-14

IF NO EFFORT MADE - SKIP TO QUESTION 9

(If yes - Probes)

a. Please describe what has been done. When?

b. What was the role of the board in the overall process from start to finish?

c. What role did you have in the process?

d. What role did your CEO/ED have in the evaluation process?

7. Have the results of the evaluation been used internally and/or externally?

(If yes - Probes)

a. Beyond publication

8. Who funded the evaluation?

9. Is there a line item in your budget for program evaluation?

(If yes - Probes)

a. Who decided to budget for the evaluation?

b. Was evaluation budgeted for in advance of the evaluation?

- c. Was the budget for evaluation a one time occurrence or does your overall budget have an ongoing line item in it for evaluation?
(If no - Probe)
 - a. Are there plans to budget for program evaluation? Regularly or annually?

10. Does your program/s have a logic model?

11. Does your organization have a current strategic plan?
(If yes - Probes)
 - a. Is program evaluation included in that plan?
 - b. If so, who suggested it be in the plan? Why?(If no - Probe)
 - a. Are there plans to develop a strategic plan?

IF NO EFFORT MADE, SKIP TO QUESTION 13

12. Do you feel the evaluation effort was beneficial to your organization?
(Probes)
 - a. If yes, why?
 - b. What were the most important benefits for you as board chair?
 - c. If no, why not?

13. Are you satisfied with how your organization is currently engaged in evaluating its program(s)?

If so, why?

Why not? If not, what do you suggest be done to increase the level at which your organization is engaged in evaluating its programs?

14. What in your opinion are the greatest challenges for evaluating your organization's program(s)?

IF EFFORT MADE SKIP TO QUESTION 16

15. Could you discuss some of the reasons why your programs have not been evaluated to date?

(Probes)

a. Resources – staff, expertise, funding, time?

16. Do you have any future plans to evaluate your program/s?

(If yes - Probe)

a. Do you expect program evaluation to benefit your organization in any way?

17. Is program evaluation related to your organization's mission?

(If yes/no - Probes)

a. If so, how?

b. If not, why not?

18. Tenure as Board member _____ YEARS _____ MONTHS

19. Top 3 priorities for the Board

20. Top 3 priorities for you as Board Chair

21. How many Board members currently serve on your Board?

22. How often does your Board meet?

23. Is the board's performance evaluated?

e. If so, how?

f. If so, how often?

g. If not, why not?

24. Have you received any instruction in evaluation? Yes_____ No_____

If yes Workshops_____ Seminars_____ College courses_____

Other _____

25. What is your highest level of education?

High School _____ Some College _____

Bachelor's Degree _____ Some graduate School _____

Master's Degree _____ Professional Degree (MD, JD, etc.) _____

Ph.D. _____

26. What was your major area of study in college or graduate school?

Public Administration _____ Social Work _____ Law _____

Business _____ Education _____ Social Sciences _____

Humanities _____ Other (specify) _____

Appendix E

FUNDING ORGANIZATION SURVEY

Name: _____ Organization: _____

1. What comes to mind when you hear the term *program evaluation*.

2. What types of information about the programs you fund do you ask from your funded organizations to report back to you after they are awarded funding? (ANSWER YES OR NO)

program activities	<input type="checkbox"/>
program finances/budget	<input type="checkbox"/>
program evaluation	<input type="checkbox"/>
Other	

3. If you selected “program evaluation,” what type/s of information do you ask from your funded organizations? (ANSWER YES OR NO)

program logic model	<input type="checkbox"/>
program process	<input type="checkbox"/>
number of people served	<input type="checkbox"/>
service delivery “best practices”	<input type="checkbox"/>
program outcomes	<input type="checkbox"/>

demonstration of use of evaluation results ☐

Other

4. Please indicate the level at which you factor in each item into your decision to continue funding programs. (REVIEW LEVELS. PLEASE SELECT ONLY ONE (1) LEVEL FOR EACH ITEM)

LEVELS

<u>ITEMS</u>	<u>Not at All</u>	<u>Somewhat</u>	<u>Strongly</u>	<u>N/A</u>
program logic model	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
number of people served	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
program outcomes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
program process evaluation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
service delivery “best practices”	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
demonstration of use of evaluation results	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
fulfill grant requirements	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Other

5. Do you include funds to pay for the evaluation of your funded program/s?

YES ☐ NO ☐ SOMETIMES ☐

Probe: If the organization asked for p.e. funds would you consider providing them?

6. If “YES”, frequently does your organization fund the evaluation of your funded program/s? (PLEASE INDICATE PERCENTAGES) If “NO,” please skip to question 10.

Include evaluation funds in the grant and
funded organization is responsible for evaluation _____

Hire and pay an outside evaluator _____

Use our own employed evaluator _____

Other _____

7. How does your organization determine the amount of funds allocated for evaluating your funded program/s?

Pre-determined amount regardless of program ☐

Percentage of total funding amount for program ☐

Driven by program parameters (delivery, outcomes) ☐

Other ☐

If “Other”, please indicate factor/s for determining amount.

I don’t know ☐

8. To what extent does your Board of Directors support the effort to have your funded programs evaluated? (PLEASE CHECK ONLY ONE)

Not at all ☐

They somewhat support it ☐

They make it a priority ☐

I don't know

☐

9. Is the evaluation of your funded programs related to your organization's mission?

Probe: If so, how?

10. Do funders similar to your type (community foundation, private foundation, UW, etc.) currently adequately fund program evaluation?

Probe: If yes, why?

If no, why not? What can be done to increase funding for p.e.?

11. Have you ever received instruction in evaluation? YES ☐ NO ☐

12. If "YES," what type of instruction have you received? (PLEASE CHECK ALL THAT APPLY)

½ day workshop ☐

1 day workshop ☐

2-5 days of instruction ☐

College or university course/s ☐

Other ☐

If "Other", please indicate the type of instruction you have received.

13. What is your highest level of education?

High School _____ Some College _____

Bachelor's Degree _____ Some graduate School _____

Master's Degree _____ Professional Degree (MD, JD, etc.) _____

Ph.D. _____

14. What was your major area of study in college or graduate school?

Public Administration _____ Social Work _____ Law _____

Business _____ Education _____ Social Sciences _____
 Humanities _____ Other (specify) _____

15. What is the size of your organization's current annual budget for funding nonprofit programs?

\$0-\$499,999	<input type="checkbox"/>	\$500,000-\$999,999	<input type="checkbox"/>
\$1-\$5 million	<input type="checkbox"/>	\$5-\$10 million	<input type="checkbox"/>
\$10-\$25 million	<input type="checkbox"/>	\$25-\$50 million	<input type="checkbox"/>
\$50-\$100 million	<input type="checkbox"/>	\$100-\$200 million	<input type="checkbox"/>
\$200 million+	<input type="checkbox"/>		

16. Do you have anything to add about program evaluation we haven't covered?

Appendix F

Interview Instrument for Evaluators of nonprofit human service programs

1. What comes to mind when you hear the term *program evaluation*?
2. What have been some of the ways in which you were selected as the evaluator?
(Probe)
 - a. Who made the final decision to select you?
3. Who has paid you for your services?
(Probes)
 - a. The organization running the program?
 - b. The funder of the program?
 - c. Other?
4. How would you define *evaluation capacity*?
5. Do you utilize logic models when you evaluate nonprofit human service programs?
(If yes - Probes)
 - a. How often?
 - b. Do logic models have a role in evaluation capacity building?
6. How often through your experience has program evaluation been a part of a nonprofit human services organization's strategic plan?
(Probe) For those that it was, did this contribute to evaluation capacity building?
7. What role does an evaluator have in helping an organization build *evaluation capacity* for its programs?

8. Have you played this role with any of the organizations that run the programs you have evaluated?
(If yes - Probes)
 - a. What specific actions have you taken?
 - b. What evaluation capacity building tools or methods have you used?
9. Have you played this role specifically with nonprofit human service organizations that run the programs you have evaluated?
(If yes - Probes)
 - a. What specific actions have you taken?
 - b. What evaluation capacity building tools or methods have you used?
10. When you think about evaluating a nonprofit human service program, who else has a role in evaluation capacity building?
11. What can these people specifically do to help build evaluation capacity?
12. How important is the relationship between you and the stakeholders involved in the program?
 - a. The management of the nonprofit human services organization?
 - b. The board?
 - c. Program staff?
 - d. Any other stakeholders?
13. What aspects of these relationships help in building evaluation capacity?
14. What are the biggest obstacles to building evaluation capacity within a nonprofit human services organization?
 - a. Anything specific to NHSOs as opposed to arts for example?
15. What do you suggest that should be done to address these obstacles?
16. Based on your experience (added) what have been the biggest challenges for you as an evaluator when dealing with evaluation capacity in nonprofit human services organizations?

17. What do you suggest be done to address (replaces “remove”)these challenges?
18. What has contributed to successful evaluation capacity building within your work with nonprofit human service programs?
(Probes)
- a. Organizational context (culture, leadership, hierarchical support, etc.)
 - b. Resources (financial, human, supplies, etc.)
 - c. Structures (policies, procedures, data collection systems, etc.)
 - d. Any recommended practices from the profession? (added)
19. How long have you worked as an evaluator? _____
20. How large is your firm?
h. Number of employees?
21. What have been the types of organizations for whom you have evaluated their programs?
- i. Human Services
 - j. Arts
 - k. Education
 - l. Health
 - m. Other
22. What has been the average compensation you have received for your evaluation work?
23. Have you received any instruction in evaluation?
(If yes - Probe) – Workshops, seminars, college courses?
24. Highest Level of Education High School_____ Some College_____
- Bachelor’s Degree_____ Some graduate School_____Master’s
- Degree_____
- Professional Degree (MD, JD, MPA, MBA, MSW) _____ Ph.D._____
25. Major area of study in college or graduate school:
- Public Administration _____ Social Work _____ Law _____
- Business _____ Education _____ Social Sciences _____ Humanities _____
- Other (specify)_____

Appendix G

Interview Instrument for Program Staff of Nonprofit Human Service Organizations

Name _____ **Organization**

1. What comes to mind when you hear the term *program evaluation*?

2. Have there been efforts made to evaluate your program/s?

IF EFFORT MADE – CONTINUE QUESTIONS 2-11

IF NO EFFORT MADE - SKIP TO QUESTION 12

(If yes - Probes)

a. Please describe what has been done. When? How often?

c. What methods were used?

d. Who conducted the evaluation?

a. Who selected the evaluator?

e. What was your role in the overall process from start to finish?

f. Does your executive director have a role in the evaluation process?

g. Does your board have a role in the evaluation process?

h. Do the consumers of the program(s) have a role in the evaluation process?

3. Have the evaluation results been used internally and/or externally?

(If yes - Probes)

a. Beyond publication

b. Program alteration – reduction, expansion, improvement

4. What costs are involved with the evaluation process?

(Probes)

- a. Did the cost include staff time allocated for the evaluation?
 - b. Were there any trade-off costs for your time spent with the evaluation process?
- 5. Who funds the evaluation process?
(Probes)
 - a. Do you feel the financial support for evaluation is adequate enough for your organization to evaluate your programs? Why? Why not?
 - b. If not, can anything be done about it?
- 6. Does your organization have plans to continue evaluating its programs?
(Probe)
 - a. If so, how often?
- 7. What are the greatest challenges for evaluating your programs?
(Probe)
 - a. What are your greatest challenges for involvement in the evaluation process?
 - b. Any other observations you have of the evaluation processes you have experienced?
- 8. If you had a choice, would you have your organization continue to have its programs evaluated?
(Probes)
 - a. If yes, why?
 - b. If no, why not?
- 9. Do you feel the evaluation effort has been beneficial to your organization?
(Probes)
 - a. If yes, why?
 - b. Were there and benefits for you as program staff?
 - c. If no, why not?
- 10. Does your program/s have a logic model?
(Probes)

- a. If so, did you participate in its development?
- b. Has the logic model been helpful in any way?

11. Have you had a role in the development and/or planning of your programs?

(Probe)

- a. If so, please describe your involvement.

IF EFFORT MADE, SKIP TO QUESTION 14

12. Could you discuss some of the reasons why your programs have not been evaluated to date?

(Probes)

- a. Resources – staff, expertise, funding, time?

13. Does your organization have any future plans to evaluate your program/s?

(If yes - Probe)

- a. Do you expect program evaluation to benefit your organization in any way?

14. Is program evaluation related to your mission?

(Probes)

- a. If so, how?
- b. If not, why not?

15. Tenure as program staff _____

16. Top 3 priorities as program staff

17. Have you received any instruction in evaluation?

(If yes - Probe) – Workshops, seminars, college courses?

18. Highest Level of Education High School _____ Some College _____
 Bachelor's Degree _____ Some graduate School _____ Master's
 Degree _____
 Professional Degree (MD, JD, MPA, MBA, MSW) _____ Ph.D. _____

19. Major area of study in college or graduate school:

Public Administration _____ Social Work _____ Law _____
 Business _____ Education _____ Social Sciences _____ Humanities

Other (specify) _____

References

- Abma, T.A. (2006). The Practice and Politics of Responsive Evaluation. *American Journal of Evaluation*, 27 (1), 31-43.
- American Camping Association. (2008). *Program Design & Activities*. Retrieved May 8, 2008 from <http://www.acacamps.org/knowledge/program/>.
- American Camping Association. (2008). *Mission & Outcomes*. Retrieved May 8, 2008 from <http://www.acacamps.org/knowledge/mission/>.
- American Counseling Association (ACA). (2008). *ACA Code of Ethics*. Retrieved May 8, 2008 from <http://www.counseling.org/Resources/CodeOfEthics/TP/Home/CT2.aspx>.
- American Evaluation Association. (2008). *About Us*. Retrieved May 8, 2008 from <http://www.eval.org/aboutus/organization/aboutus.asp>.
- American Evaluation Association. (2008). *Find an Evaluator*. Retrieved May 8, 2008 from http://www.eval.org/find_an_evaluator/evaluator_search.asp.
- American Evaluation Association. (2004). *Guiding Principles for Evaluators*. Retrieved March 8, 2008 from <http://www.eval.org/Publications/GuidingPrinciples.asp>.
- American Evaluation Association. (2008). *Local Affiliates*. Retrieved May 8, 2008 from <http://www.eval.org/aboutus/organization/affiliates.asp>.
- American Evaluation Association. (2008). *Summer Evaluation Institute*. Retrieved May 8, 2008 from <http://www.eval.org/SummerInstitute08/08SIhome.asp>.
- American Evaluation Association. (2008). *University Programs*. Retrieved May 8, 2008 from http://www.eval.org/Training/university_programs.asp.
- American Psychological Association (APA). (2002). *Ethical Principles of Psychologists and Code Of Conduct*. Retrieved April 30, 2008 from <http://www.apa.org/ethics/code2002.pdf>.
- Amo, C., & Cousins, J.B. (2007). GoingThrough the Process: An Examination of the Operationalization of Process Use in Empirical Research on Evaluation. *New Directions for Evaluation*, 116, 5-26.
- Andresen, E.M., Machuga, C.R., Van Booven, M., Egel, J., Chibnall, J.T., & Tait, R.C. (2008). Effects and Costs of Tracking Strategies on Nonresponse Bias in a Survey of Workers with Low-Back Injury. *Public Opinion Quarterly*, 72 (1), 40-54.

- Anthes, E.W. (1987). *Personnel Matters in the Nonprofit Organization*. Hampton, AR: Independent Community Consultants, Inc.
- Argyris, C., & Schon, D.A. (1996). *Organizational Learning II: Theory, Method, and Practice (2nd Edition)*. Reading, MA: Addison-Wesley Publishing Co.
- Axelrod, N.R. (2005). Board Leadership and Development. In R.D. Herman & Associates (Eds.), *The Jossey-Bass Handbook of Nonprofit Leadership and Management* (pp. 131-152). San Francisco: John Wiley & Sons.
- Babbie, E. (1995). *The Practice of Social Research*. Belmont, CA: Wadsworth Publishing Company.
- Baizerman, M., Compton, D.W., & Stockdill, S.H. (2002). New Directions for ECB. *New Directions for Evaluation*, 93, 109-119.
- Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real World Evaluation*. Thousand Oaks, CA: Sage Publications, Inc.
- Barker, R.L. (1999). *Social Work Dictionary* (4th edition). Washington, DC: NASW Press.
- Barnette, J.J., & Wallis, A.B. (2003). Helping Evaluators Swim with the Current: Training Evaluators to Support Mainstreaming. *New Directions for Evaluation*, 99, 51-61.
- Beatty, P.C. (2004). The Dynamics of Cognitive Interviewing. In S. Presser, J.M. Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, J. Martin & E. Singer, (Eds.), *Methods for Testing and Evaluating Survey Questionnaires* (pp. 45-66). Hoboken, NJ: John Wiley & Sons, Inc.
- Beatty, P.C., & Willis, G.B. (2007). Research Synthesis: The Practice of Cognitive Interviewing. *Public Opinion Quarterly*, 71 (2), 287-311.
- Bell, J.B. (1994). Managing Evaluation Projects Step by Step. In J.H. Wholey, H.P. Hatry and K.E. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (pp. 510-533). San Francisco: Jossey-Bass Publishers.
- Bennis, W. (1989). *Why leaders Can't Lead*. San Francisco: Jossey-Bass, Inc.
- Berger, P., & Luckman, T. (1967). *The Social Construction of Reality*. New York: Anchor Books.
- Bernstein, D.J., Whitsett, M.D., & Mohan, R. (2002). Addressing Sponsor and Stakeholder Needs in the Evaluation Authorizing Environment: Trends and Implications. *New Directions for Evaluation*, 95, 89-99.

- Bjerke, B. (1999). *Business Leadership and Culture*. Northampton, MA: Edward Elgar Publishing, Ltd.
- Block, S.R. (2001a). Board of Directors. In J. Steven Ott (Ed.), *Understanding Nonprofit Organizations*, (pp. 15-24). Boulder, CO: Westview Press.
- Block, S.R. (2001b). Executive Director. In J. Steven Ott, Ed. *Understanding Nonprofit Organizations*, (pp. 100-107). Boulder, CO: Westview Press.
- Block, S.R. (2004). *Why Nonprofits Fail*. San Francisco: Jossey-Bass.
- Bolman, L.G., & Deal, T. (2003). *Reframing Organizations: Artistry, Choice and Leadership*. San Francisco: John Wiley & Sons.
- Boulmetis, J., & Dutwin, P. (2005). *The ABCs of Evaluation*. San Francisco: Jossey-Bass.
- Boyle, R., & Lemaire, D. (1999). *Building Effective Evaluation Capacity: Lessons from Practice*. New Brunswick, NJ: Transaction Publishers.
- Bresnick, D.A. (1983). *Managing the Human Services in Hard Times*. New York: Human Services Press.
- Brinckerhoff, P. (2000). *Mission-Based Management*. New York City: John Wiley & Sons.
- Brody, E. (2002). Accountability and Public Trust. In L. Salamon (Ed.), *The State of Nonprofit America* (pp. 471-498). Washington, DC: Brookings Institution Press.
- Brun, C.F. (2005). *A Practical Guide to Social Service Evaluation*. Chicago: Lyceum Books.
- Bryson, J.M. (2005). The Strategy Change Cycle: An Effective Strategic Planning Approach for Nonprofit Organizations. In R.D. Herman & Associates (Eds.), *The Jossey-Bass Handbook of Nonprofit Leadership and Management* (pp. 171-203). San Francisco: John Wiley & Sons.
- Burgess, B. (1993). The Board of Directors. In T.D. Connors (Ed.), *The Nonprofit Management Handbook: Operating Policies and Procedures* (pp. 195-227). New York: John Wiley & Sons.
- Campbell, David. (2002). Outcomes Assessment and the Paradox of Nonprofit Accountability. *Nonprofit Management and Leadership*, 12 (3), 243-259.
- Carlson, M., & Donohoe, M. (2003). *The Executive Director's Survival Guide*. San Francisco: John Wiley & Sons.

- Carman, J.G. (2005). *Program Evaluation Use and Practice: A Study of Nonprofit Organizations in New York State* (Doctoral Dissertation, University at Albany, State University of New York, 2005).
- Carman, J.G. (2007). Evaluation Practice Among Community-Based Organizations. *American Journal of Evaluation*, 28 (1), 60-75.
- Carver, J. (2002). *On Board Leadership*. San Francisco: Jossey-Bass.
- Centers for Disease Control and Prevention (CDC). (1997). *Framework for Program Evaluation in Public Health*. Retrieved March 6, 2008 from <http://www.cdc.gov/mmwr/preview/mmwrhtml/rr4811a1.htm>.
- Chen, H.T. (2005). *Practical Program Evaluation*. Thousand Oaks, CA: Sage Publications.
- Chickering, A.W. (1977). *Experience and Learning: An Introduction to Experiential Learning*. New Rochelle, NY: Change Magazine Press.
- Christensen, R.A., & Ebrahim, A. (2006). How does accountability affect mission? The case of a nonprofit serving immigrants and refugees. *Nonprofit Management and Leadership*, 17 (2), 195-209.
- Cimmino, P.F. (2004). Basic Concepts and Definitions of Human Services. In H.S. Harris, D.C. Maloney & F.M. Rother (Eds.), *Human services: Contemporary Issues and Trends* (pp. 5-17). Boston: Allyn and Bacon.
- Compton, D.W., Glover-Kudon, R., Smith, I.E., & Avery, M.E. (2002). Ongoing Capacity Building in the American Cancer Society. *New Directions for Evaluation* 93, 47-62.
- Compton, D.W., & Baizerman, M. (2007). Defining Evaluation Capacity Building. *American Journal of Evaluation* 28 (1), 118-119.
- Cooke, R.A., & Szumal, J.L. (2006). Using the Organizational Culture Inventory to Understand the Operating Cultures of Organizations. *Handbook of Organizational Culture and Climate* (pp. 147-162). Thousand Oaks, CA: Sage Publishing.
- Cousins, J.B. & Whitmore, E. (1998). Framing Participatory Evaluation. *New Directions for Evaluation* 80, 87-105.
- Davis, S. (1984). *Managing Corporate Culture*. Cambridge, MA: Ballinger Publishing Company.
- Decatur Cooperative Ministry. (2007). *Mission and Vision*. Retrieved April 1, 2007 from <http://www.decaturcooperativeministry.org/missionvision2.htm>.

- Denhardt, R.B., Denhardt, V.J., & Aristigueta, M.P. (2002). *Managing Human Behavior in Public & Nonprofit Organizations*. Thousand Oaks, CA: Sage Publications, Inc.
- DeRobertis, E.M., & Saldarini, R. (2004). Technology and Sociality in the New Millennium: Current Challenges for the Human Service Generalist. In H.S. Harris, D.C. Maloney & F.M. Rother (Eds.), *Human services: Contemporary Issues and Trends* (pp. 375-381). Boston: Allyn and Bacon.
- Diambra, J.F. (2004). Historical Roots of Human Services. In H.S. Harris, D.C. Maloney & F.M. Rother (Eds.), *Human services: Contemporary Issues and Trends* (pp. 23-40). Boston: Allyn and Bacon.
- DiMaggio, P., & Powell, W.W. (1991). Introduction. In W.W. Powell and P.J. DiMaggio (Eds.), *The New Institutionalism in Organizational Analysis* (pp. 1-38). Chicago: University of Chicago Press.
- Donovan, L.L., Meyer, S.R., & Fitzgerald, S.P. (2007). Transformative Learning and Appreciative Inquiry: A More Deeper Union for Organizational Change. *Academy of Management Proceedings* pp. 1-6.
- Drucker, P.F. (1967). *The Effective Executive*. New York: Harper & Row.
- Drucker, P.F. (1990). *Managing the Nonprofit Organization*. New York: Harper Collins Publishers.
- Duigan, P. (2003). Mainstreaming Evaluation or Building Evaluation Capability? Three Key Elements. *New Directions for Evaluation*. 99, 7-21.
- Dym, B., & Hutson, H. (2005). *Leadership in Nonprofit Organizations*. Thousand Oaks, CA: Sage Publications.
- Eliason, M.J. (2007). *Improving Substance Abuse Treatment*. Thousand Oaks, CA: Sage Publications, Inc.
- Elleven, R.K. (2007). Appreciative Inquiry: A Model for Organizational Development and Performance Improvement in Student Affairs. *Education* 127 (4), 451-455.
- Eriksen, K. (1977). *Human Services Today*. Reston, VA: Reston Publishing Company.
- Etzioni, A. (1964). *Modern Organizations*. Englewood Cliffs, NJ: Prentice Hall.
- Families First. (2007). *Welcome to Families First*. Retrieved April 1, 2007 from <http://www.familiesfirst.org/index.html>.

- Ferris, J. (1993). The Double-Edged Sword of Social Service Contracting: public accountability versus nonprofit autonomy. *Nonprofit Management and Leadership*, 3 (4), 363-376.
- Festen, M. & Philbin, M. (2007). *Level Best: How Small and Grassroots Nonprofits can Tackle Evaluation and Talk Results*. San Francisco: Jossey-Bass.
- Fetterman, D.M. (2003). Fetterman-House: A Process Use Distinction and a Theory. *New Directions for Evaluation* 97, 47-52.
- Fetterman, D. M. (2005). Conceptualizing Empowerment in Terms of Sequential Time and Social Space. In D,M, Fetterman & A. Wandersman (Eds.), *Empowerment Evaluation Principles in Practice* (pp. 209-214). New York: The Guilford Press.
- Fetterman, D. M. (2005). Empowerment Evaluation Principles in Practice. In D,M, Fetterman & A. Wandersman (Eds.), *Empowerment Evaluation Principles in Practice* (pp. 42-72). New York: The Guilford Press.
- Fine, A.H., Thayer, C.E., & Coghlan, A.T. (2000). Program Evaluation Practice in the Nonprofit Sector. *Nonprofit Management and Leadership*, 10 (3), 331-339.
- Fitzpatrick, J.L., Sanders, J.R. & Worthen, B.R. (2004). *Program Evaluation: Alternative Approaches and Practical Guides*. Boston: Pearson Education.
- Fredericks, K.A., Carman, J.G., & Birkland, T.A. (2002). Program Evaluation in a Challenging Authorizing Environment: Intergovernmental and Interorganizational Factors. *New Directions for Evaluation* 95, pp. 5-21.
- Frost, P.J., Moore, L.F., Louis, M.R., Lundberg, C.C., & Martin, J. (1985). An Allegorical View of Organizational Culture. In P.J. Frost, L.F. Moore, M.R. Louis, C.C. Lundberg, & J. Martin (Eds.), *Organizational Culture* (pp. 13-25). Beverly Hills, CA: Sage Publications.
- Gann, N. (2001). Toward a Contract Culture. In J. Steven Ott, Ed. *Understanding Nonprofit Organizations*, (pp. 247-255). Boulder, CO: Westview Press.
- Gardner, F. (2006). *Working With Human Service Organisations*. New York: Oxford University Press.
- Gibelman, M. (2003). *Navigating Human Service Organizations*. Chicago: Lyceum Press, Inc.
- Gray, S.T., & Stockdill, S.H. (1995). *Leadership IS: Evaluation With Power*. Washington, DC: Independent Sector.

- Green, F.L. (2004). Ten Things Nonprofits Must Do in the Twenty-First Century. In R.E. Riggio, & S.S. Orr (Eds.), *Improving Leadership in Nonprofit Organizations* (pp. 19-35). San Francisco: John Wiley & Sons.
- Grinnell, Jr., R.M. & Unrau, Y.A. (2005). *Social Work Research and Evaluation*. New York: Oxford University Press.
- Groves, R.M., Fowler, Jr., F.J., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2004). *Survey Methodology*. Hoboken, NJ: John Wiley & Sons.
- Grudens-Schuck, N. (2003). The Rigidity and Comfort of Habits: A Cultural and Philosophical Analysis of the Ups and Downs of Mainstreaming Evaluation. *New Directions for Evaluation*, 99, 23-32.
- Guba, E.G. & Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage Publications, Inc.
- GuideStar. Retrieved 9/4/2005 from <http://www.guidestar.org>.
- Gurin, A., & Friedman, B. (1989). The Efficacy of Contracting for Service. In H.W. Demone, Jr., & M. Gibelman (Eds.), *Services for Sale: Purchasing Health and Human Services*. New Brunswick, NJ: Rutgers University Press.
- Guzman, B.L. & Feria, A. (2002). Community-Based Organizations and State Initiatives: The Negotiation Process of Program Evaluation. *New Directions for Evaluation*, 95, pp. 57-72.
- Hawkins, B.O. (1984). The Role of Evaluation in Philanthropic Decision Making. (Doctoral dissertation, University of Minnesota, 1984). *UMI Dissertation Abstracts*, 8424701.
- Hay, R.D. (1990). *Strategic Management in Nonprofit Organizations*. Westport, CT: Greenwood Press.
- Herman, R.D., & Heimovics, D. (1991). *Executive Leadership in Nonprofit Organizations: New Strategies for Shaping Executive-Board Dynamics*. San Francisco: Jossey-Bass.
- Herman, R.D., & Heimovics, D. (2005). Executive Leadership. In R.D. Herman & Associates (Eds.), *The Jossey-Bass Handbook of Nonprofit Leadership and Management* (pp.153-170). San Francisco: John Wiley & Sons.
- Herman, R.D., & Renz, D.O. (2002). *Nonprofit Organizational Effectiveness: Practical Implications Of Research On An Elusive Concept*. Midwest Center for Nonprofit Leadership - Research. Retrieved April 9, 2008 from <http://bsbpa.umkc.edu/mwcnl/research/RESEARCH.HTM>.

- Hoefler, R. (2000). Accountability in action?: Program evaluation in nonprofit human service agencies. *Nonprofit Management and Leadership*, 11 (2), 167-177.
- Holbrook, A., Cho, Y.I., & Johnson, T. (2006). The Impact of Question and Respondent Characteristics on Comprehension and Mapping Difficulties. *Public Opinion Quarterly* 70 (4), pp. 565-595.
- Holland, T.P. (1995). Organizations: Context for social service delivery. In R.L. Edwards (Ed.) *Encyclopedia of Social Work* (19th ed., pp. 1787-1794). Washington, DC: NASW Press.
- Holstein, J.A., & Gubrium, J.F. (1995). *The Active Interview*. Thousand Oaks, CA: Sage Publishing.
- Hudson, M. (2005). *Managing at the Leading Edge*. San Francisco: Jossey-Bass.
- Hull, G.H., Jr. & Kirst-Ashman, K.K. (2004). *The Generalist Model of Human Services Practice*. Pacific Grove, CA: Thomson Brooks/Cole.
- Ivancevich, J.M., Szilagyi, Jr., A.D., & Wallace, Jr., M.J. (1977). *Organizational Behavior and Performance*. Santa Monica, CA: Goodyear Publishing Company, Inc.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage. "Citing the Standards" Retrieved May 8, 2008 from Western Michigan University at <http://www.wmich.edu/evalctr/jc>.
- Kaplan, R.S., & Norton, D.P. (2006). *Alignment: Using the Balanced Scorecard to Create Corporate Strategies*. Boston: Harvard Business School Press.
- Keeter, S., Kennedy, C., Dimock, M., Best, J., & Craighill, P. (2006). Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey. *Public Opinion Quarterly* 70 (5), 759-779.
- King, J.A. (2004). Tikkun Olam: The Roots of Participatory Evaluation. In M.C. Alkin (Ed.), *Evaluation Roots*, (pp. 331-342). Thousand Oaks, CA: Sage Publications, Inc.
- King, J.A. (2005). A Proposal to Build Evaluation Capacity at the Bunche-DaVinci Learning Partnership Academy. *New Directions for Evaluation* 106, 85-97.
- King, J.A. (2007). Developing Evaluation Capacity Through Process Use. *New Directions for Evaluation* 116, 45-59.

- Kraines, G. (2001). *Accountability Leadership: How to Strengthen Productivity Through Sound Managerial Leadership*. Franklin Lakes, NJ: Career Press.
- Leighninger, Leslie (ed.). (2002). Social Work Training Needs: Yesterday and Today. *Journal of Progressive Human Services*, 13 (2), 61-65.
- Lipsey, M.W. (1993). Theory as Method: Small Theories of Treatments. *New Directions for Program Evaluation* 57, 30-62.
- Lynch, R. (1993). *Lead! How Public and Nonprofit Managers can Bring out the Best in Themselves and their Organizations*. San Francisco: Jossey-Bass.
- Malloy, C.L., & Yee, P.A. (2006). Client Relations: More than just "Business." *New Directions for Evaluation* 111, 67-77.
- Marais, L.C. (1998). The Relationship Between organizational Culture and the Practice of Program Evaluation in Human Service Organizations (Doctoral dissertation, , Western Michigan University, 1998). UMI Dissertation Services, 9840039.
- Mark, M.M., Henry, G.T., & Julnes, G. (2000). *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Public and Nonprofit Policies and Programs*. San Francisco: Jossey-Bass.
- Martin, L., & Kettner, P.M. (1996). *Measuring the Performance of Human Service Programs*. Thousand Oaks, CA: Sage Publications, Inc.
- Mattesich, P.W. (2003). *The Manager's Guide to Program Evaluation*. St. Paul: Wilder Publishing Center.
- Mehr, J.J., & Kanwischer, R. (2004). *Human Services: Concepts and Intervention Strategies*. Boston: Pearson Education, Inc.
- Merriam-Webster. (2005). *Webster's Third New International Dictionary, Unabridged*. Springfield, MA: Merriam-Webster, Inc.
- Mesch, D.J., & McClelland, J. (2006). Managing for Performance and Integrity. In D.R. Young (Ed.), *Wise Decision-Making in Uncertain Times: Using Nonprofit Resources Effectively*. New York: Foundation Center.
- Miller, T.I., Kobayashi, M.M., & Noble, P.M. (2006). Insourcing, not Capacity Building, a Better Model for Sustained Program Evaluation. *American Journal of Evaluation* 27 (1), 83-94.
- Milstein, B., Chapel, T.J., Wetterhall, S.F., & Cotton, D.A. (2002). Building Capacity for Program Evaluation at the Centers for Disease Control and Prevention. *New Directions for Evaluation* 93, 27-46.

- Minkoff, D.C. & Powell, W.W. (2006). Nonprofit Mission: Constancy, Responsiveness, or Deflection? In W.W. Powell & R. Steinberg (Eds.) *The Nonprofit Sector: A Research Handbook* (pp.591-611). New Haven: Yale University Press.
- Mohan, R., & Sullivan, K. (2006). Managing the Politics of Evaluation to Achieve Impact. *New Directions for Evaluation* 112, 7-23.
- Monette, D.R., Sullivan, T.J., & DeJong, C.R. (2002). *Applied Social Research: Tool for the Human Services*. Orlando: Harcourt Publishers.
- Monroe, M.C., Fleming, M.L., Bowman, R.A. Zimmer, J.F., Marcinkowski, T. Washburn, J. et al. (2005). Evaluators as Educators: Articulating Program Theory and Building Evaluation Capacity. *New Directions for Evaluation* 108, 57-71.
- Mothers Against Drunk Driving (MADD). (2005). *25 Years of Saving Lives, 1980-2005*. Retrieved May 8, 2008 from <http://www.madd.org/getattachment/48e81e1b-df43-4f31-b9a1-d94d5b940e62/MADD-25-Years-of-Saving-Lives.aspx>.
- Murray, V. (2005). Evaluating the Effectiveness of Nonprofit Organizations. In R.D. Herman & Associates (Eds.), *The Jossey-Bass Handbook of Nonprofit Leadership and Management* (pp.345-370). San Francisco: John Wiley & Sons.
- Nanus, B., & Dobbs, S.M. (1999). *Leaders who Make a Difference: Essential Strategies for Meeting the Nonprofit Challenge*. San Francisco: Jossey-Bass.
- National Association of Social Workers. (1999). *Code of Ethics*. Retrieved May 10, 2006, from <http://www.socialworkers.org/pubs/code/default.asp>.
- National Association of Social Workers. (2007). *Evidence-Based Practice*. Retrieved May 13, 2008 from <http://www.socialworkers.org/research/naswResearch/0108EvidenceBased/default.asp>.
- National Association of Social Workers. (2005). *NASW Standards for Social Work Practice in Child Welfare*. Retrieved March 26, 2008, from <http://www.socialworkers.org/practice/standards/NASWChildWelfareStandards0905.pdf>.
- National Association of Workforce Development Professionals (NAWDP). (2002). *NAWDP Code of Professional Ethics and Practices*. Retrieved April 30, 2008 from <http://dev56.vandammehosting.com/Content/NavigationMenu/About/CodeofEthics/ethics.pdf>.

- National Center for Charitable Statistics (NCCS). (2007). *Number of Public Charities in the United States, 2006*. Retrieved May 13, 2008 from <http://nccsdataweb.urban.org.proxy.ulib.iupui.edu/PubApps/profileDrillDown.php?state=US&rpt=PC>.
- National Organization for Human Service Education (NOHSE) & Council for Standards in Human Service Education (CSHSE). (2004). The Human Service Worker: A Generic Job Description. In H.S. Harris, D.C. Maloney & F.M. Rother (Eds.), *Human services: Contemporary Issues and Trends* (pp. 123-128). Boston: Allyn and Bacon.
- Newcomer, K.E., Hatry, H.P., & Wholey, J.S. (1994). Meeting the Need for Practical Evaluation Approaches: An Introduction. In Wholey, Joseph H., Harry P. Hatry and Kathryn E. Newcomer, (Eds.), *Handbook of Practical Program Evaluation* (pp. 1-10). San Francisco: Jossey-Bass Publishers.
- Nonprofit and Foundation Evaluation American Evaluation Association TIG. (2006). *About Us – Membership*. Retrieved May 8, 2008 from http://www.nonprofitandfoundationtig.org/index_files/Page347.html.
- Nord, W.R. (1985). Can Organizational Culture be Managed? In P.J. Frost, L.F. Moore, M.R. Louis, C.C. Lundberg, & J. Martin (EDs.), *Organizational Culture* (pp. 187-196). Beverly Hills, CA: Sage Publications.
- Northouse, P. (2004). *Leadership: Theory and Practice*. Thousand Oaks, CA: Sage Publications, Inc.
- O'Connell, B. (1985). *The Board Members Book*. New York: The Foundation Center.
- Olson, K. & Peytchev, A. (2007). Effect of Interviewer Experience on Interview Pace and Interviewer Attitudes. *Public Opinion Quarterly* 71(2), 273-286.
- Oster, S.M. (1995). *Strategic Management for Nonprofit Organizations: Theory and Cases*. New York: Oxford University Press.
- Ostroff, C., & Schmitt, N. (1993). Configurations of Organizational Effectiveness and Efficiency. *Academy of Management Journal* 36 (6), 1345-1361.
- Ott, J.S. (1989). *The Organizational Culture Perspective*. Chicago: The Dorsey Press.
- Paddock, S.C. (2001). Evaluation. In J.S. Ott (Ed.), *Understanding Nonprofit Organizations: Governance, Leadership, and Management* (pp. 359-365). Boulder, CO: Westview Press.
- Palumbo, D.J. (1987). Politics and Evaluation. In D.J. Palumbo (Ed.), *The Politics of Program Evaluation* (pp. 12-46). Beverly Hills, CA: Sage Publications.

- Patton, M.Q. (1997). *Utilization-Focused Evaluation: The New Century Text*. Thousand Oaks, CA: Sage Publishing.
- Patton, M.Q., Bare, J. & Bonnet, D.G. (2004). Building Strong Foundation-Grantee Relationships. In M.T. Braverman, N.A. Constantine, & J.K. Slater (Eds.), *Foundations and Evaluation: Contexts and Practices for Effective Philanthropy* (pp. 76-95). San Francisco: Jossey-Bass.
- Pawlak, E.J., & Vinter, R.D. (2004). *Designing and Planning Programs for Nonprofit & Government Organizations*. San Francisco: Jossey-Bass.
- Pfeffer, J., & Salancik, G.R. (1978). *The External Control of Organizations: Resource Dependence Perspective*. New York: Harper & Row.
- Pfeffer, J., & Salancik, G.R. (2003). *The External Control of Organizations: A Resource Dependence Perspective*. Stanford, CA; Stanford University Press.
- Plantz, M.C., Greenway, M.T. & Hendricks, M. (1997). Outcome Measurement: Showing Results in the Nonprofit Sector. *New Directions for Evaluation* 75, 15-30.
- Poister, T.H. (2003). *Measuring Performance in Public and Nonprofit Organizations*. San Francisco: Jossey-Bass.
- Polkinghorne, D.E. (2004). *Practice and the Human Sciences*. Albany: State University of New York Press.
- Poole, D.L., Davis, J.K., Reisman, J., & Nelson, J.E. (2001). Improving the Quality of Outcome Evaluation Plans. *Nonprofit Management and Leadership*, 11 (4), 405-421.
- Powell, J.L. (1995). *Pathways to Leadership: How to Achieve and Sustain Success*. San Francisco: Jossey-Bass.
- Presser, S., Rothgeb, J.M., Couper, M.P., Lesser, J.T., Martin, E., Martin, J., & Singer, E. (Eds.), (2004). *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken, NJ: John Wiley & Sons.
- Preskill, H. (2004). The Transformational Power of Evaluation. In M.C. Alkin (Ed.), *Evaluation Roots* (pp. 343-355). Thousand Oaks, CA: Sage Publications.
- Presskill, H., & Catsambas, T.T. (2006). *Reframing Evaluation Through Appreciative Inquiry*. Thousand Oaks, CA: Sage Publishing.
- Presskill, H., & Russ-Eft, D. (2005). *Building Evaluation Capacity*. Thousand Oaks, CA: Sage Publications.

- Presskill, H., & Torres, R.T. (1999). *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, CA: Sage Publications.
- Reamer, F.G. (1998). *Social Work and Evaluation Skills*. New York: Columbia University Press.
- Reingold, D. (2006). Are Independent Program Evaluations Really Independent? Unpublished paper.
- Rogers, P.J., Petrosino, A., Huebner, T.A., & Hacs, T.A. (2000). Program Theory Evaluation: Practice, Promise and Problems. *New Directions for Evaluation* 87, 5-13.
- Rooney, P. & Frederick, H.K. (2007). *Paying for Overhead: A Study of the Impact of Foundations' Overhead Payment Policies on Educational and Human Service Organizations*. Retrieved July 5, 2008 from <http://www.philanthropy.iupui.edu/Research/WorkingPapers/PayingforOverhead.pdf>.
- Rothman, R.A. (1987). *Working: Sociological Perspectives*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Royse, D. (1991). *Research Methods in Social Work*. Chicago: Nelson-Hall Publishers.
- Rubin, A. (2006). Forward. *The Research Process in the Human Services*. Belmont, CA: Thompson Books/Cole.
- Rubin, H. J. & Rubin, I.S. (2005). *Qualitative Interviewing: The Art of Hearing Data*. Thousand Oaks, CA: Sage Publications, Inc.
- Russ-Eft, D. & Presskill, H. (2001). *Evaluation in Organizations*. New York: Basic Books.
- Salamon, L. (2002). The Resilient Sector. In L. Salamon (Ed.), *The State of Nonprofit America* (pp. 3-61). Washington, DC: Brookings Institution Press.
- Sanders, J.R. (2003). Mainstreaming Evaluation. *New Directions for Evaluation*, 99, 3-6.
- Schalock, Robert L. (1995). *Outcome-Based Evaluation*. New York: Plenum Press.
- Schein, E. (1992). *Organizational Culture and Leadership*. San Francisco: Jossey-Bass Publishers.

- Scheirer, M. A. (1994). Designing and Using Process Evaluation. In J.H. Wholey, H.P. Hatry, & K.E. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (pp. 40-68). San Francisco: Jossey-Bass Publishers.
- Schuerman, J.R. (1983). *Research and Evaluation in the Human Services*. New York: The Free Press.
- Schumpeter, J.A. (1991). *The Economics and Sociology of Capitalism*. Princeton, NJ: Princeton University Press.
- Schwandt, D.R., & Marquardt, M.J. (2000). *Organizational Learning: From World-Class Theories to Global Best Practices*. New York: St. Lucie Press.
- Scribner, S.M. (2004). Last Things First – Theories and Realities: A Perspective on Nonprofit Leadership. In R.E. Riggio, & S.S. Orr (Eds.), *Improving Leadership in Nonprofit Organizations* (pp. 131-147). San Francisco: Jossey-Bass.
- Scriven, M. (1993). Hard-Won Lessons in Program Evaluation. *New Directions for Evaluation* 58. San Francisco: Jossey-Bass Publishers.
- Search Institute. (2007). *Why Are the 40 Developmental Assets Important?* Retrieved May 10, 2008 from <http://www.search-institute.org/assets/importance.html>.
- Senge, P.M. (1990). *The Fifth Discipline*. New York: Currency Doubleday.
- Senge, P.M. (1996). Leading Learning Organizations. In F. Hesselbein, M. Goldsmith, & R. Beckhard (Eds.), *The Leader of the Future* (pp. 41-57). San Francisco: Jossey-Bass.
- Sheppard, M. (2004). *Appraising and Using Social Research in the Human Services*. Philadelphia: Jessica Kingsley Publishers.
- Simon, H.A. (1997). *Models of Bounded Rationality, Vol. 3: Empirically Grounded Economic Reason*. Cambridge, MA: The MIT Press.
- Smith, Bucklin & Associates, Inc. (Eds.). (2000). *The Complete Guide to Nonprofit Management, 2nd edition*. New York: John Wiley & Sons.
- Somers, C. (2005). Evaluation of the Wonders in Nature – Wonders in Neighborhoods Conservation Education Program: Stakeholders Gone Wild! *New Directions for Evaluation*, 108, 29-46.
- Sonnichsen, R.C. (1994). Evaluators as Chang Agents. In J.H. Wholey, H.P. Hatry, & K.E. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (pp. 534-548). San Francisco: Jossey-Bass Publishers.

- Sonnichsen, R.C. (1999). Building Evaluation Capacity Within Organizations. In R. Boyle, & D. Lemaire (Eds.), *Building Effective Evaluation Capacity: Lessons from Practice*. New Brunswick, NJ: Transaction Publishers.
- Stake, R.E. (2001). A Problematic Heading. *American Journal of Evaluation*, 22 (3), 349-354.
- Stern, G.J. (1999). *The Drucker Foundation Self-Assessment Tool: Process Guide*. New York: The Peter F. Drucker Foundation for Nonprofit Management.
- Stevahn, L., King, J.A., Ghore, G., & Minnema, J. (2005). Establishing Essential Competencies for Program Evaluators. *American Journal of Evaluation*, 26 (1), 43-59.
- Stinchcombe, A. (1965). Social Structure and Organizations. In J. G. March (Ed.), *Handbook of Organizations* (p. 142-193). Chicago: Rand-McNally.
- Stockdill, S.H., Baizerman, M. & Compton, D.W. (2002). Toward a Definition of the ECB Process: A Conversation with the ECB Literature. *New Directions for Evaluation*, 93, 7-26.
- Stufflebeam, D. (2002). *Institutionalizing Evaluation Checklist*. The Evaluation Center: Western Michigan University. Retrieved March 18, 2007 from: <http://www.wmich.edu/evalctr/checklists/institutionalizingeval.pdf>.
- Stufflebeam, D. (2004). The 21st-Century CIPP Model. In M.C. Alkin (Ed.), *Evaluation Roots* (pp. 245-266). Thousand Oaks, CA: Sage Publications, Inc.
- Swidler, A. (1986). Culture in Action: Symbols and Strategies. *American Sociological Review*, 51 (2), 273-86.
- Tang, H., Cowling, D.W., Koumjian, K., Roesler, A., Lloyd, J. & Rogers, T. (2002). Building Local Program Evaluation Capacity Toward a Comprehensive Evaluation. *New Directions for Evaluation*, 95, 39-56.
- The Evaluators' Institute. (2008). *Overview of Certificates*. Retrieved May 9, 2008 from http://www.evaluatorsinstitute.com/certificate_program_overview.php.
- Thomas, J.C. (2005). Outcome Assessment and Program Evaluation. In R.D. Herman & Associates (Eds.), *The Jossey-Bass Handbook of Nonprofit Leadership and Management* (pp. 391-416). San Francisco: John Wiley & Sons.
- Thompson, N. (2000). *Theory and Practice in Human Services*. Philadelphia: Open University Press.

- Torres, R.T. & Preskill, H. (2001). Evaluation and Organizational Learning: Past, Present and Future. *American Journal of Evaluation*, 22 (3), 387-395.
- Trevisan, M.S. (2007). Evaluability Assessment from 1986 to 2006. *American Journal of Evaluation*, 28 (3), 290-303.
- Tripodi, T. (1983). *Evaluative Research for Social Workers*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Tschirhart, M. (1996). *Artful Leadership: Managing Stakeholder problems in Nonprofit Arts Organizations*. Bloomington, IN: Indiana University Press.
- Underwood, R. & Lee, M. (2004). Research Models. In H.S. Harris, D.C. Maloney & F.M. Rother (Eds.), *Human Services: Contemporary Issues and Trends* (pp. 387-405). Boston: Allyn and Bacon.
- United Way of America. (1996). *Measuring Program Outcomes: A Practical Approach*. Alexandria, VA: United Way of America.
- Van Slyke, D. (2003). The Mythology of Privatization in Contracting for Social Services. *Public Administration Review*, 63 (3), 296-315.
- Volkov, B., & King, J. (2005). A Grounded Checklist for Implementing Evaluation Capacity Building in Organizations. Unpublished paper presented at the joint meeting of the American Evaluation Association and the Canadian Evaluation Society, Toronto, November, 2005.
- Volkov, B & King, J. (2007). *A Checklist for Building Organizational Evaluation Capacity*. The Evaluation Center: Western Michigan University. Accessed March 28, 2008 from <http://www.wmich.edu/evalctr/checklists/ecb.pdf>.
- Wandersman, A., Snell-Johns, J., Lentz, B.E., Fetterman, D.M., Keener, D.C., Livet, M., et al. (2005). The Principles of Empowerment Evaluation. In D.M. Feterman & A. Wandersman (Eds.), *Empowerment Evaluation Principles in Practice* (pp. 27-41). New York: The Guilford Press.
- Warren, R.L. (1974). The Social Context of Evaluation Research. In W.C. Sze & J.G. Hopps (Eds.), *Evaluation and Accountability in Human Service Programs* pp. 13-35. Cambridge, MA: Schenkman Publishing.
- Waterman, Jr., R.H., Peters, T.J., & Phillips, J.R. (1980). Structure is not Organization. *Business Horizons*, 23 (3), 14-25.
- Weber, M. (1949). Objectivity in the Social Sciences. In E. Shils (Ed), *Max Weber on the Methodology of the Social Sciences*, Glencoe, IL: The Free Press.

- Weiss, C. (1998). *Evaluation*, 2nd edition. Upper Saddle River, NJ: Prentice Hall.
- Weiss, J. A. (2000). From Research to Social Improvement: Understanding Theories of Intervention. *Nonprofit and Voluntary Sector Quarterly*, 29 (1), 81-110.
- Werther, Jr., W.B. & Berman., E.M. (2001). *Third Sector Management: The Art of Managing Nonprofit Organizations*. Washington, DC: Georgetown University Press.
- Wholey, J.H. (1994). Assessing the Feasibility and Likely Usefulness of Evaluation. In J.H. Wholey, H.P. Hatry & K.E. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (pp. 15-39). San Francisco: Jossey-Bass Publishers.
- Wiley, J.W., & Brooks, S.M. (2006). The High-Performance Organizational Climate. *Handbook of Organizational Culture and Climate*, pp.177-191. Thousand Oaks, CA: Sage Publishing.
- Williams, D.D. & Hawkes, M.L. (2003). Issues and Practices Related to Mainstreaming Evaluation: Where Do We Flow from Here? *New Directions for Evaluation*, 99, 63-83.
- Williamson, O.E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: The Free Press.
- W.K. Kellogg Foundation. (1998). *W.K. Kellogg Foundation Evaluation Handbook*. Retrieved May 8, 2008 from <http://www.wkkf.org/default.aspx?tabid=101&CID=281&CatID=281&ItemID=2810770&NID=20&LanguageID=0>.

¹ See Charity Navigator at <http://www.charitynavigator.org/>.

² See Better Business Bureau's Wise Giving Alliance at <http://us.bbb.org/WWWRoot/SitePage.aspx?site=113&id=4ef08b14-37cb-4974-a385-7f41f63b16b0>.

³ Joint Standards can be accessed at <http://www.eval.org/EvaluationDocuments/progeval.html>.

⁴ Guiding Principles for Evaluation can be accessed at <http://www.eval.org/Publications/GuidingPrinciples.asp>

⁵ American Evaluation Association – <http://www.eval.org>.

⁶ See Russ-Eft & Preskill's *Evaluation in Organizations*, p. 421 for complete instrument.

⁷ For more complete, detailed coverage of ECB, see the following resources – Stufflebeam's *Institutionalizing Evaluation Checklist* at <http://www.wmich.edu/evalctr/checklists/institutionalizingeval.pdf>; Volkov and King's *Checklist for Building Organizational Evaluation Capacity* at <http://www.wmich.edu/evalctr/checklists/ecb.pdf>; Baizerman, Compton and Stockdill's volume "The Art, Craft, and Science of Evaluation Capacity Building" in *New Directions for Evaluation*, Number 93, Spring, 2002 (especially page 111), and Preskill and Russ-Eft's *Building Evaluation Capacity* (2005).

⁸ See Stevahn, L., King, J.A., Ghore, G., & Minnema, J. (2005). Establishing Essential Competencies for Program Evaluators. *American Journal of Evaluation* 26(1), pp. 49-51 for complete taxonomy.

⁹ See standards 1-8 for "Standards for Minimal Disclosure" in American Association for Public Opinion Research's (AAPOR) *Code of Professional Ethics and Practices* adopted 2005, published 2006. Accessible at <http://www.aapor.org/aaporcodeofethics>.

¹⁰ GuideStar – <http://www.guidestar.org>.

¹¹ GuideStar changed its income level categories during the research for this study, however the original categories were maintained for consistency between the two MSA samples. With 21 organizations and five income levels in each MSA sample, an extra organization would be the fifth in a given income level category. The Atlanta sample had 5 organizations in the \$1,000,000-\$4,999,999 income level category while the Indy sample had 5 organizations in the \$250,000-\$499,999 income level category.

¹² Organizations either refusing to participate, out of business, unable to be reached or incorrectly listed in the human services category by GuideStar were replaced with other randomly selected organizations in that income level and county.

¹³ NVIVO is a trademark of QSR International.

¹⁴ Note: Variance and bias exist in this rating process, as not all 42 organizations had their ED, Board Chair and program staff representative interviewed. Some had all three, some had two of the three and some had only one stakeholder interviewed. Therefore, the organizations' ratings were based on data from however many interviews were conducted from that organization's stakeholders, as well as the available documentation provided by that organization.

Curriculum Vitae

Salvatore Alaimo

Education

Ph.D., Philanthropic Studies – Indiana University – 2008

Doctoral Minor: Nonprofit Management

Dissertation: Program Evaluation Capacity for Nonprofit Human Service Organizations: An Analysis of Determining Factors

M.S., Urban Studies/Nonprofit Management – Georgia State University – 1999

B.S., Business Administration – Ramapo State College of New Jersey – 1988

Concentration: Operations Management

Training

Certified in Volunteer Administration (CVA)

Graduate of United Way's V.I.P. board development program

Graduate of Leadership DeKalb (DeKalb County, GA)

Professional Experience

9/05 – present Adjunct Professor – Indiana University-Purdue University Indianapolis

5/98 – present Consultant for nonprofit organizations

7/02 – 8/04 Adjunct Professor – Georgia State University

2/97 – 6/04 Outcomes Measurement Manager – Girl Scout Council of NW Georgia

1/96 – 1/97 Program Coordinator – United Way of Metropolitan Atlanta

Publications

Nonprofits and Evaluation: Managing Expectations from the Leader's Perspective – in *New Directions for Evaluation* 119, edited by Kim Fredericks and Joanne Carman, Jossey-Bass

Contracting Out – forthcoming chapter in *Nonprofit Economics and Management*, edited by Bruce Seaman and Dennis Young, Edward Elgar Publishing

Service Learning Program Evaluation – Briefing Paper for *Learning to Give*, <http://www.learningtogive.org/papers/index.asp?bpid=178>

Highlights of Giving for the Past 50 Years – *Giving USA 2005 50th Anniversary* (Contributor and research assistant)

Program outcomes study featured in *Measuring Performance in Public and Nonprofit Organizations* by Theodore H. Poister, San Francisco: Jossey-Bass, 2003, pp. 231-233

Research Presentations

Qualitative Research Panel. Association for Research on Nonprofit Organizations and Voluntary Action Annual Conference, Philadelphia (November, 2008)

Corporate Social Responsibility: Questions Raised About Partnering with Employee Volunteer Programs. Association of Volunteer Resources Management 2nd annual conference, Binghamton, NY (October, 2008)

Nonprofits and Evaluation: Managing Expectations from the Leader's Perspective. American Evaluation Association's 2007 Conference, Baltimore (November, 2007)

Corporate Community Outreach in Metro Atlanta. The Foundation Center, Atlanta (April, 2007)

Capacity for Program Evaluation in Nonprofit Human Service Organizations: An Analysis of Determining Factors. Andrew Young School of Policy Studies at Georgia State University, Atlanta (February, 2007)

Examining the Role of the Leader in Organizational Capacity for Program Evaluation. Association for Research on Nonprofit Organizations and Voluntary Action Annual Conference, Chicago (November, 2006)

The Role of Leadership in Organizational Capacity for Program Evaluation. American Evaluation Association's 2006 Conference, Portland (November, 2006)

The Role of Leadership in Organizational Capacity for Program Evaluation. Campbell Public Affairs Institute, The Maxwell School at Syracuse University (October, 2005)

Program Evaluation in Nonprofit Human Service Organizations: A Conceptual Theory of Ethical Responsibility. Association for Research on Nonprofit Organizations and Voluntary Action Annual Conference, Washington, DC (November, 2005)

Nonprofits and the Services they Contract for: Who, What for and How much? Association for Research on Nonprofit Organizations and Voluntary Action Annual Conference, Washington, DC (November, 2005)

The Role of Leadership in Organizational Capacity for Program Evaluation. Association for Research on Nonprofit Organizations and Voluntary Action Annual Conference, Washington, DC (November, 2005)

Courses taught

PAUS 8921A – Volunteer Management (graduate); summer 2002 & 2004 – Georgia State University

SPEA V-362 – Nonprofit Management and Leadership (undergraduate); fall 2005 & spring 2006 – IUPUI

SPEA V-525 – Management in the Nonprofit Sector (graduate); fall 2007 & 2008 – IUPUI

Manuscript Reviews

Nonprofit and Voluntary Sector Quarterly 2006-2008
The Jossey-Bass Handbook of Nonprofit Leadership & Management (3rd edition) 2008
Evaluation and Program Planning 2008

Workshop/Panel Presentations

Strategic Planning for Implementation. Enterprise Community Partners, Enterprise Foundation

Diversity in Philanthropy: What is the Relationship to Effectiveness in Grantmaking? Council on Foundations, Foundation Center, Association for Research on Nonprofit Organizations & Voluntary Action (ARNOVA) Researcher/Practitioner Forum

Using Your Program Data. Enterprise Community Partners, Enterprise Foundation

Designing Programs for Measuring Effectiveness.” The Foundation Center
“Building Successful Partnerships: The Keys to Succeeding with Consultants.” The Nonprofit Risk Management Center

Risk Management in Your Volunteer Program. Georgia Conference on Service & Volunteerism

Success with Volunteers. Georgia Special Olympics

Finders, Keepers: How to Recruit and Retain Volunteers. Prevent Child Abuse Georgia

Evaluating your Volunteer Program. Central Indiana Association for Volunteer Administration

Evaluating your Volunteer Program. Central Indiana Corporate Volunteerism Council

Building Capacity for Program Evaluation. United Way of Metro Atlanta V.I.P. Alumni Association

The Evaluation Component of Your Grant Proposal. The Foundation Center

Evaluating Adult Education. Georgia State University

Measuring your Employee Volunteer Program. Metro Atlanta Corporate Volunteer Council

Honors, Awards, Fellowships

Dissertation research grants – The Center on Philanthropy, 2006 & 2008
Lumina Foundation for Education, 2006

Educational Enhancement Grant – IUPUI Graduate Student Organization, 2007

Doctoral scholarship – IUPUI, 2006-2007

Doctoral fellowship – IUPUI, 2004-2006

ARNOVA Conference Scholarship – 2006

Board Experience

President Elect – Atlanta Area Evaluation Association

President – Council of Volunteer Administrators for Metro Atlanta

Research and Evaluation Advisory Board – Prevent Child Abuse Georgia

Fund Development Committee – Starfish Initiative

Service

Public Relations Committee – Association of Volunteer Resource Managers conference

Program Assessment Committee – Decatur Cooperative Ministry

Technical advisor to United Way agencies – Atlanta Area Evaluation Association

References Committee Chair – Association for Volunteer Administration

Strategic planning committee co-chair – United Way VIP Alumni Association

Professional Memberships

Academy of Management

American Association of Public Opinion Research (AAPOR)

American Evaluation Association (AEA)

Association for Research on Nonprofit Organizations and Voluntary Action (ARNOVA)